ReCALL (2019), 31: 2, 116–134
doi:10.1017/S0958344018000083

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Developing oral communication in Spanish lower-level courses: The case of voice recording and videoconferencing activities

Marta Tecedor
Arizona State University, USA (marta.tecedor@asu.edu)

Gonzalo Campos-Dintrans
University of Mary Washington, USA (gcamposd@umw.edu)

Abstract
This study adds to the literature on computer-mediated communication (CMC) by examining the impact of online voice recording (VR) activities and peer-to-peer videoconferencing (VC) conversations on the development of beginning Spanish learners' speaking performance. Specifically, this paper explores (1) whether VR and VC activities promote oral proficiency, and if so, whether those gains can be seen both in presentational and interpersonal modes of communication; and (2) whether VR and VC activities foster oral proficiency in similar ways to face-to-face (F2F) communication. A quasi-experimental, pre-/post-test design was used in the study. Three sections of a first semester Spanish course were assigned to one of the following conditions: F2F, VR, and VC. Complexity, accuracy, and fluency measures were used to analyze learners' speaking performance in the two tasks. A mixed effects model analysis was used to investigate differences across time as well as among groups. Results show that both F2F communication and VR activities promote complexity and fluency in presentational tasks and fluency in interpersonal tasks, although F2F produces superior results regarding complexity in presentational tasks. VC activities promote complexity and fluency in presentational tasks and complexity, accuracy, and fluency in interpersonal tasks. Overall, this study shows that medium is not merely a delivery device but has important implications for learning outcomes. In this sense, these findings contribute to answering the wider question of how the use of technology in second language instruction plays a decisive role in current teaching practices.

Keywords: CMC tools; voice recording; videoconferencing; CAF; beginning learners

1. Introduction

As oral communication becomes an increasingly important goal in foreign/second language education (Sadoux, 2013), the use of computer-mediated voice-based tools to help students improve their speaking abilities has experienced a rise in popularity, as attested by the numerous textbooks that now include these tools. The assumption behind this trend is that providing students with opportunities to practice speaking via computer-mediated-communication (CMC) tools helps them improve their oral proficiency. However, we still have limited empirical evidence outlining how exactly these new forms of technology help beginning learners improve their

Cite this article: Tecedor, M. & Campos-Dintrans, G. (2019). Developing oral communication in Spanish lower-level courses: The case of voice recording and videoconferencing activities. *ReCALL* 31(2): 116–134, doi: 10.1017/S0958344018000083

oral communication skills – data that is particularly important for classroom instructors, given that not all CMC tools enable the same tasks and learner actions.

Currently, two of the most common technology-mediated instructional practices for the development of second language (L2) oral skills are voice-recording (VR) activities, in which speakers produce language without immediate interaction, and peer-to-peer videoconferencing conversation (VC) activities, in which speakers interact via technology in a synchronous manner. Although these activities share many characteristics, there are also important differences between them. VR elicits brief samples of presentational language typical, for example, of phone voice-mail, in which the speaker records a message for an interlocutor who cannot respond in real time. VC, on the other hand, can be used to elicit interpersonal language typical of casual conversations with friends and family.

Common beliefs about these two types of activities, in part pushed by the publishing houses, are that VR activities promote oral proficiency and help beginning learners prepare for face-to-face (F2F) communication,[1] whereas VC activities are thought to mimic F2F interaction, and as such, are assumed to be as conducive to L2 acquisition as in-person communication (Wang, 2004). However, without empirical evidence to support these beliefs, the accuracy of these statements needs to be taken with caution. From a theoretical standpoint, examining the potential positive effect of VR and VC activities is also warranted, given that there is sufficient evidence in the literature to hypothesize that VR and VC are conducive to language learning. Namely, the main tenets of the Output Hypothesis (Swain, 1985, 1995) suggest that giving learners the opportunity to orally express complete thoughts in the target language translates into an improved ability to converse, which is provided by both VR and VC activities. VC activities are also thought to be facilitative, taking into consideration the Interaction Hypothesis (Long, 1996), which states that negotiated interaction helps learners achieve a more target-like output. Thus, there is enough theoretical grounding to examine whether and how these two types of activities benefit L2 learners, and there is also an empirical motivation to do so, given how widespread the alleged benefits are.

Previous studies that have examined the development of language abilities through the use of voice-based CMC tools (Guillén & Blake, 2017; Sykes, 2005; Xiao, 2007; Yanguas, 2010) have shown that voice-based tools can have a positive impact on L2 speaking development. However, none of them have addressed the question of whether such development affects interpersonal and presentational modes of communication alike. This study aims to fill this gap by examining (1) whether VR activities in fact promote oral proficiency, and if so, whether those gains can be seen both in presentational and interpersonal modes of communication; and (2) whether VC activities foster oral proficiency in similar ways to F2F communication, and if so, whether those gains can also be extended to presentational language. Thus, this study contributes to this field of research by providing further empirical evidence of the connections between voice-based CMC and the development of the aforementioned modes of communication in terms of complexity, accuracy, and fluency, and as such, it is of special interest to L2 practitioners and researchers alike.

2. Literature review

2.1 Technology and L2 speaking development

The field of computer assisted language learning (CALL) has had a long tradition of exploring how technology affects the development of reading (Chun, 2006, 2011), writing (Elola & Oskoz, 2011; Murray & Hourigan, 2006), and listening (Jones, 2006; Robin, 2011). In the area of

[1]Previous to the publication of this article, Pearson Education advertised in its platform: "Oral Recording Activities enable students to increase oral proficiency without leaving the online environment." Vista Higher Learning similarly stated: "Language production can be intimidating for new students, but this initial step helps build confidence and better prepare students for face-to-face communication."

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

speaking, however, there was initially more skepticism with regard to the potential effectiveness of technology. The focus of CALL research in this area was initially in pronunciation and intonation training (O'Brien, 2006). With the advent of CMC tools in the early 1990s, a cohort of studies began examining the transferability of language skills from synchronous written tools (i.e. chat) to oral production (e.g. Abrams, 2003; Beauvois, 1998). A second wave of research in the early 2000s aimed at examining new CMC voice-based tools, such as voice recordings, voice blogs, voice threads, and videoconferencing.

CMC voice-based tools can be broadly classified into two different groups. Asynchronous tools, such as VoiceThread, Wimba Voice, and audio and video blogs, allow for the transmission of information in an intermittent manner. Thus, they represent a one-way communication genre that elicits presentational language. The main advantages of these types of tools are that they allow planning time and sometimes the possibility of multiple recordings. In contrast, synchronous tools, such as audio- and videoconferencing platforms (e.g. Skype, Google Hangouts), favor real-time two-way exchanges in which interlocutors have to constantly and actively negotiate the meaning and make adjustments in order to achieve their communicative goals. Synchronous tools, then, work best for interpersonal communication. The incorporation of audio and visual features in these tools simulates the sense of immediacy characteristic of F2F communication, which has led experts to depict videoconferencing as the most interactive CMC-based environment to date (Wang, 2004) and as "the best medium for interaction where face to face contact is not feasible" (Goddard, 1995: 213).

In its relatively short history, the literature in voice-based CMC tools has taken on a number of different perspectives. The majority of studies have examined its affective benefits, in particular participants' attitudes towards the tools per se (Bueno Alastuey, 2011; Wang, 2006), their potential for reducing language anxiety (Poza, 2005; Cho & Carey, 2001), and their ability to increase L2 motivation (Gleason & Suvorov, 2012) and self-confidence (Kern, 2014). Although these studies are useful to understand how L2 learners respond to CMC interactions, they provide only anecdotal support to the claim that voice-based CMC tools help improve the quality of spoken discourse.

Compared to the affective domain, fewer studies have addressed the development of formal linguistic features through the use of voice-based CMC tools, although those that do exist agree that such tools have a positive influence on L2 speaking development. In one of the first studies aimed at comparing the impact of different interactional environments, Sykes (2005) investigated the acquisition of refusal speech acts by third semester Spanish L2 learners in written chat, oral chat, and F2F exchanges. The results of both quantitative and qualitative analyses suggested that the writing chat group outperformed both the oral chat and the F2F groups in terms of the complexity and the variety of refusals produced, but the oral chat group displayed more varied types of strategies for managing the interaction.

In a similar study, Yanguas (2010) measured the effects of three synchronous interactional environments during a jigsaw-style task. In particular, he compared how third semester Spanish L2 learners negotiated for meaning in audioconferencing, videoconferencing, and F2F conversational exchanges. The results of the analysis showed no differences between the videoconferencing and the F2F groups. Furthermore, the participants in the audioconferencing group presented more instances of negotiation of meaning than those in the videoconferencing and the F2F groups, probably due to the lack of visual channel.

In further support of voice-based CMC tools for the development of L2 oral proficiency, Xiao (2007) compared linguistic gains – measured in terms of complexity, accuracy, and fluency – by L2 learners of English interacting through videoconferencing and in F2F conversations over the course of 10 weeks. The results of the quantitative analysis showed a significant improvement in fluency (syllables per minute) and accuracy (error-free clauses), but not in complexity (T-units per minute) for L2 learners in the videoconferencing group. Additionally, a positive interaction was found between the videoconferencing experience and increased oral achievement. That is,

participants in the videoconferencing group outperformed those in the F2F group in all measures. However, the difference in interlocutors – native speakers for the experimental group versus non-native speakers for the control group – makes it difficult to attribute these results solely to the effect of the interactional mode.

More recently, Guillén and Blake (2017) examined the effect of asynchronous video recordings and synchronous videoconferencing tandem exchanges in two Spanish intermediate hybrid courses. Results of the analysis showed that asynchronous video recordings encouraged the use of complex syntactic structures to levels similar to those of written discourse. The results also showed that asynchronous video recordings produced longer T-units than the tandem exchanges. The authors explain these results in terms of the affordances provided by the medium. The asynchronous tool, due to the additional planning time it provides, allows students to produce more elaborate constructions, whereas the synchronous tool fosters fluency development measured in terms of sentence mastery, which is a strong indicator of language automatization and therefore it is helpful in developing L2 speaking abilities.

In summary, previous research has supported the idea that voice-based CMC tools have a positive impact on L2 speaking. The following section looks at the affordances of the two voice-based CMC tools under study – VR and VC – in light of second language acquisition (SLA) theories of L2 development.

2.2 L2 development

From an SLA theoretical standpoint, CMC voice-based tools' effectiveness can be explained in terms of two hypotheses: the Output Hypothesis (Swain, 1985, 1995) and the Interaction Hypothesis (Long, 1996).

The main tenet of the Output Hypothesis is that L2 learners need to engage in form in addition to meaning in order to fully process grammatical structures. Swain (1985) suggested that output is one of the ways to push students' focus on these two aspects of the message they want to convey – moving from semantic to syntactic processing – and to reformulate their own utterances (modified output). In addition, producing language can help learners (a) notice the gaps in their interlanguage so they can pay attention to the linguistic features they lack in future input opportunities; (b) provide opportunities to test whether their use of the L2 is effective and/or accurate; (c) develop metalinguistic awareness; and (d) develop fluency as a by-product of practice (Gass & Mackey, 2006). In the current study, both VR and VC activities provide opportunities to produce output and consequently they may both help learners engage in form as well as meaning.

In addition to output, L2 acquisition has also been linked to receiving feedback and negotiating meaning. According to Long's (1996) Interaction Hypothesis, acquisition occurs when L2 learners modify their output in response to a communication breakdown or to feedback from their interlocutor. Because of this, conversational interactions are a valuable learning experience in that they allow learners to not only put into practice what they have previously learned, but also to hear and produce language in ways that can go beyond their existing knowledge.

The acquisitional benefits of both output and interaction have been found in numerous studies (e.g. Gass & Mackey, 2006; Izumi, 2003). However, one of the main topics of contention in the literature regards the types of tasks that lead to greater acquisition. Although one-way tasks provide greater opportunities for L2 learners to produce modified output than two-way tasks (e.g. Iwashita, 1999; Shehadeh, 1999), two-way tasks seem to promote greater accuracy and long-term retention rates of grammatical forms (e.g. Swain, 1997). The current study seeks to add to the discussion of the impact of task features in L2 development by examining how presentational tasks – in this case, via VR activities – and presentational tasks – in this case, via VC activities – fulfill output and interactional needs of the L2 learning process.

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

2.3 Measuring L2 speaking performance

L2 speaking production has traditionally been evaluated according to three measures: complexity, accuracy, and fluency (CAF). These measures have formed the foundation of L2 speaking evaluations since the 1970s, and yet many questions remain regarding their employment in instructional settings, including how CAF should be defined and operationalized and to what extent these measures relate and/or interact during L2 performance and development.

Complexity, commonly defined as the ability to use a wide and varied range of structures, has been operationalized in terms of formal (e.g. range of word types, number of word families) and functional (e.g. number of words, clauses, or subordinate clauses per T-, C-, or AS-unit) features. *Accuracy* is characterized as the ability to produce target-like language; that is, error-free production. It has been measured both in terms of specific measures (e.g. percentage of target-like verbal morphology, use of plural or gender markers) and in more general terms (e.g. percentage of error-free clauses, errors per 100 words). Finally, *fluency* is defined in the literature as ease of production, and in studies of speaking performance it has been measured in terms of rate of speech (e.g. syllables per minute, number of pauses, length of pauses) or dysfluencies (e.g. number of false starts, repetitions, reformulations).

Ideally, the goal of L2 instruction is to help learners develop in these three areas as they advance in their learning process. Research from a cognitive framework, however, indicates that focusing on all of the three components at the same time may not be possible. Skehan's (1998) Limited Capacity Hypothesis states that L2 processing demands, along with learners' limited attentional resources, create a tension between the three CAF traits, which results in a trade-off effect. Learners either favor the meaning of the message, which results in greater fluency, or one of the subcomponents of form, either accuracy or complexity, but not both of them at the same time. In contrast, Robinson's (1995) Cognition Hypothesis claims that a task's characteristics (e.g. number of elements, one-way vs. two-way tasks) and administration conditions (e.g. pre-task planning time) make learners focus on either meaning or form of the message, and that the trade-off between accuracy and complexity suggested by the Limited Capacity Hypothesis may be overcome under certain circumstances. Specifically, the Cognition Hypothesis predicts that complex monologic tasks promote more accurate and more complex language to the detriment of fluency, whereas simple tasks are likely to foster only fluency.

Both hypotheses have been tested in the literature, but so far the results of empirical studies have yet to provide definitive evidence to support one over the other. There is a consensus in the literature, however, regarding trade-off effects between fluency and accuracy (e.g. Ahmadian & Tavakoli 2011; Michel, Kuiken & Vedder, 2007; Yuan & Ellis 2003). The current study, by looking at beginning learners' speaking performance, hopes to provide a clearer understanding of how trade-off effects function within the context of CMC voice-based tools.

3. Research questions

This study uses a quasi-experimental, intact class design and addresses the following questions:

1. Do VR activities help beginning learners of Spanish improve *complexity, accuracy,* and *fluency* in presentational and interpersonal modes of communication?
2. Do VC activities help beginning learners of Spanish improve *complexity, accuracy,* and *fluency* in presentational and interpersonal modes of communication?

4. Methodology

4.1 Participants and setting

The participants in this study were students enrolled in three classes of a Spanish elementary review course at a large public university in the Southwest. To enroll in this class, students had

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

either completed two years of Spanish in high school or tested into this level via the CLEP proficiency exam. Out of the initial 57 participants who provided their signed consent, nine were removed from the study for being heritage speakers of Spanish. In the end, 48 participants were retained. The classes covered chapters Preliminar to 5 of *Mosaicos*, were taught by graduate students, and followed a flipped method instruction whereby, prior to each class meeting, students completed a series of online activities, readings, and video viewings that revolved around grammatical and cultural topics. During class time, students engaged in communicative activities designed to practice targeted vocabulary and grammatical structures. Class activities were provided by the Spanish program and progressed from controlled activities that aimed at providing practice at the word and sentence levels, to free-practice activities, in which students had the opportunity to personalize the language, experiment, and incorporate previously learned language. Instructors were free to design their own class activities, but they had all received the same training on lesson plan design with both controlled and free-practice activities. Each course section met three times a week for 50 minutes, had a different instructor, and was assigned to one of the following conditions: F2F, VR, and VC. Table 1 details demographic information, providing range values, means, and standard deviations for age, years of prior Spanish instruction, academic status (years in college), and previous knowledge (chapter tests average score).

Table 1. Demographic information

	F2F group ($n = 18$)	VR group ($n = 14$)	VC group ($n = 16$)
Age	Range = 19–24	Range = 18–21	Range = 19–22
	$M = 20.8$, $SD = 1.33$	$M = 18.92$, $SD = .99$	$M = 20.25$, $SD = 1.23$
Gender	7 males, 11 females	6 males, 8 females	6 males, 10 females
Prior Spanish instruction	Range = 1–4	Range = 2–4	Range = 2–4.5
	$M = 2.5$, $SD = .92$	$M = 2.78$, $SD = .8$	$M = 2.71$, $SD = .85$
Academic status	Range = 2–5	Range = 1–3	Range = 2–4
	$M = 3.16$, $SD = .98$	$M = 1.5$, $SD = .64$	$M = 2.79$, $SD = .79$
Previous knowledge	Range = 28.5–48.5	Range = 34.5–47.75	Range = 31.75–46.25
	$M = 40.94$, $SD = 5.16$	$M = 41.34$, $SD = 4.15$	$M = 39.81$, $SD = 4.86$

4.2 Data collection

To determine the comparability of the groups in terms of linguistic knowledge prior to the completion of the experimental tasks, this study used the average score from two chapter tests administered in Weeks 5 and 7 of the semester. Each test was worth 50 points and included listening, reading, writing, vocabulary, grammar, and culture (Table 2). The results of the one-way ANOVA did not yield a significant difference between the groups, $F(2, 44) = .42$, $p = .66$, suggesting that they had comparable knowledge of Spanish and therefore were appropriate for the study.

In Weeks 8 and 15 of the semester, as pre- and post-test respectively, all groups completed a presentational and an interpersonal task. Participants' performance in the two tasks was recorded using the software Sanako. The presentational tasks consisted of a two-minute preparation time and a two-minute recording time of a monologic activity in which participants had to answer to a prompt they received in the computer (see Appendix A). During the two-minute preparation time, they were instructed to practice aloud or in silence. They could not write or talk to their

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

Table 2. Time frame of procedures

	F2F group	VR group	VC group
Week 5	Chapter test 1	Chapter test 1	Chapter test 1
Week 7	Chapter test 2	Chapter test 2	Chapter test 2
Week 8	Pre-test: presentational and interpersonal tasks	Pre-test: presentational and interpersonal tasks	Pre-test: presentational and interpersonal tasks
Weeks 9–14	Class activities	Presentational task	Interpersonal task
Week 15	Post-test: presentational and interpersonal tasks	Post-test: presentational and interpersonal tasks	Post-test: presentational and interpersonal tasks
Week 16	Participant questionnaire	Participant questionnaire	Participant questionnaire

classmates. The interpersonal task consisted of a three-minute preparation time and five-minute recording time of a dialogic, peer-to-peer role-play activity (see Appendix B). For the paired activity, participants were arranged into dyads by their instructors according to a holistic and intuitive assessment of their speaking proficiency. They remained with the same partner during the whole semester. During the three-minute preparation time they were instructed to talk to their conversation partner in Spanish or English to prepare for the task. They were not allowed to write any notes. Some practiced the language of the actual task; some focused on planning their answer.

During Weeks 9–14, the VR and the VC groups, in addition to regular classroom instruction, went to the language lab once a week (15 minutes approximately) and completed a presentational and an interpersonal activity respectively. The F2F group completed a F2F version of the interpersonal task in class. Finally, in Week 16 all participants signed the consent form and completed the participant questionnaire whose purpose was to elicit data regarding their linguistic and personal backgrounds.

4.3 Transcription and coding

Each of the presentational and interpersonal activities recorded in Weeks 8 and 15 were broadly transcribed and subsequently pruned as the total speech produced minus cut-off words, false starts, reformulations, asides in the L1, and self-corrections, in which case only the self-corrected speech was counted. Then, CAF measures were obtained by the two researchers of this study, with over 90% inter-rater reliability. The remaining discrepancies were discussed until a consensus was reached.

4.4 Measures of CAF

CAF studies differ with regard to how each of the constructs should be defined and operationalized. The purpose of this study was to use measures that were sensitive enough to detect differences between groups. Thus, following Skehan and Foster (1999), complexity was operationalized as the average number of words per AS unit. Accuracy was operationalized as the percentage of error-free AS units, the errors coded were lexical and morphosyntactic in nature, and most consisted of subject-verb and gender-number agreement in articles, quantifiers, and adjectives. Errors in pronunciation and asides in the L1 were not counted. Given the level of the participants and the distinctive nature of the two tasks under investigation, fluency was operationalized as the average number of pruned syllables per minute.

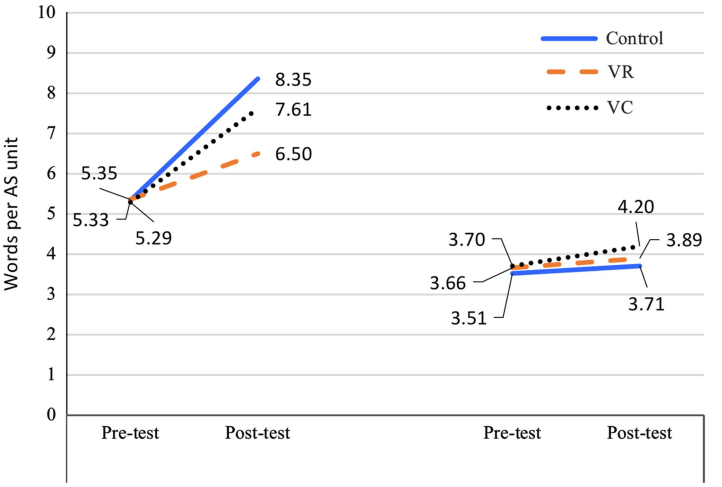
https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

5. Analysis

In order to gauge participants' performance before and after treatment, a mixed effects model analysis was carried out to compare each group's performance across time as well as among groups. In this section, we report the results of the analysis for each of the CAF components by examining (1) whether the three groups were similar at pre-test, (2) the extent to which the groups' performance changed after the treatments, and (3) the differences in gains among the three groups after the six-week intervention period.

5.1 Complexity

As previously noted, complexity was operationalized as the number of words per AS unit. Figure 1 and the accompanying table show group means and standard deviations values (in parentheses) for the pre-test and post-test, along with gain scores in both the presentational and the interpersonal tasks.



	Presentational task			Interpersonal task		
	M (*SD*)	*M* (*SD*)	Gains	*M* (*SD*)	*M* (*SD*)	Gains
F2F	5.33 (.92)	8.35 (1.22)	3.02	3.51 (.53)	3.71 (.54)	0.2
VR	5.35 (.63)	6.5 (1.49)	1.15	3.66 (.58)	3.89 (.62)	0.23
VC	5.29 (1.07)	7.61 (1.61)	2.32	3.7 (.73)	4.2 (.88)	0.5

Figure 1. Descriptive statistics of complexity

5.1.1 Presentational task

All groups performed similarly at pre-test, with around 5.3 words per AS unit. The mixed model effects test revealed no significant differences between the groups: F2F versus VC ($t = .11$, $p = .91$, $d = .39$); F2F versus VR ($t = -.07$, $p = .94$, $d = .03$); and VR versus VC ($t = -.17$, $p = .86$, $d = .74$).

The difference in mean scores at pre-test and post-test showed positive gains for all groups, and the differences were statistically significant, and with large effect sizes for all three groups: F2F ($t = 9.13$, $p = .000^*$, $d = 2.79$), VR ($t = 3.27$, $p = .001^*$, $d = 1.00$), and VC ($t = 6.17$, $p = .000^*$, $d = 1.69$).

Finally, we examined the differences in gains among the three groups, and the analysis revealed differences between the F2F and VR groups ($t = 3.88$, $p = .000^*$, $d = 1.57$), as well as between the VR and VC groups ($t = 2.27$, $p = .02^*$, $d = .97$), with large effect sizes for both comparisons. The only contrast that did not yield a statistically significant difference was between the F2F and VC groups ($t = -1.41$, $p = .15$, $d = .59$). Results of the comparisons between groups are summarized in Table 3 for ease of reference.

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

Table 3. Complexity: Group comparisons in presentational task

	VR		VC	
	Differences at pre-test	Differences in gains	Differences at pre-test	Differences in gains
F2F	$p = .94$	$p = .000^*$	$p = .91$	$p = .15$
	$d = .03$	$d = 1.57$	$d = .39$	$d = .59$
VC	$p = .86$	$p = .02^*$		
	$d = .74$	$d = .97$		

These results suggest that regardless of treatment, beginning learners improve in complexity over a six-week period. Yet they also suggest that although VR activities do not hinder development, completing VC or F2F activities has a greater impact in complexity during CMC presentational tasks.

5.1.2 Interpersonal task

Complexity scores at pre-test ranged between 3.5 and 3.7 words per AS unit. The analysis revealed no differences between the groups at pre-test: F2F versus VC ($t = -.85$, $p = .39$, $d = .30$); F2F versus VR ($t = .70$, $p = .48$, $d = .26$); and VR versus VC ($t = .17$, $p = .86$, $d = .05$).

The difference in mean scores at pre-test and post-test showed positive gains for all groups, although the gains were noticeably smaller than in the presentational task. Indeed, the F2F group's gains were not statistically significant ($t = 1.21$, $p = .22$, $d = .37$), neither were the VR group's ($t = 1.34$, $p = .18$, $d = .38$). Only the VC group's gains were statistically significant ($t = 2.67$, $p = .008^*$, $d = .62$), with a medium size effect.

Lastly, the examination of groups' gains revealed no differences between the groups: F2F versus VR ($t = -.15$, $p = .88$, $d = .54$); F2F versus VC ($t = 1.20$, $p = .22$, $d = .46$); and VC versus VR ($t = 1.04$, $p = .3$, $d = .41$), a finding at odds with the improvement observed in the VC group from the pre-test to the post-test.

Table 4. Complexity: Group comparisons in interpersonal task

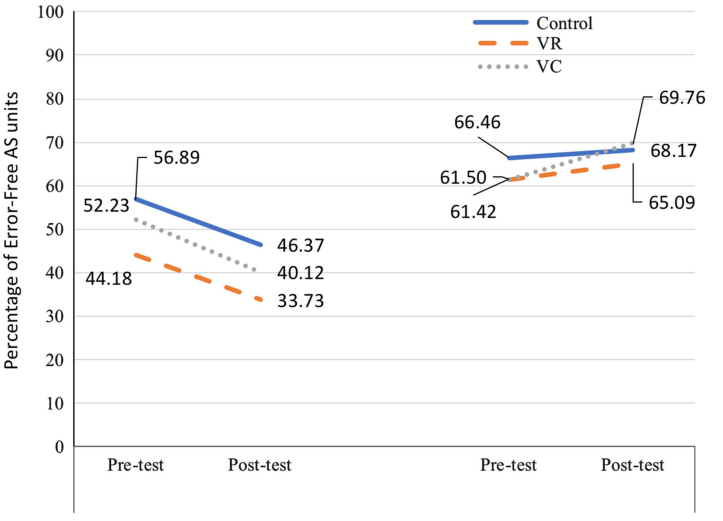
	VR		VC	
	Differences at pre-test	Differences in gains	Differences at pre-test	Differences in gains
F2F	$p = .48$	$p = .88$	$p = .39$	$p = .22$
	$d = .26$	$d = .54$	$d = .30$	$d = .46$
VC	$p = .86$	$p = .3$		
	$d = .05$	$d = .41$		

The Table 4 data suggest that completing paired VC activities has a positive impact on complexity in CMC interpersonal tasks, whereas doing VR and F2F activities has no effect. However, these results must be taken with caution, as no statistical differences were found between the VC and the other two groups when the gains scores were compared.

5.2 Accuracy

Accuracy was operationalized in terms of percentage of error-free AS units. Figure 2 and the accompanying table show group means and standard deviation values for the pre-test and post-test, along with gain scores in both the presentational and the interpersonal tasks for all groups.

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press



	Presentational task			Interpersonal task		
	M (*SD*)	*M* (*SD*)	Gains	*M* (*SD*)	*M* (*SD*)	Gains
F2F	56.89 (17.91)	46.37 (17.87)	−10.51	66.46 (9.15)	68.17 (9.54)	1.70
VR	44.18 (17.61)	33.73 (9.22)	−10.44	61.42 (11.85)	65.09 (9.07)	3.67
VC	52.23 (22.47)	40.12 (14.03)	−12.1	61.5 (10.43)	69.76 (10.32)	8.26

Figure 2. Descriptive statistics of accuracy

5.2.1 Presentational task

Accuracy scores at pre-test ranged between 44% and 56%. The results of the analysis revealed significant differences between the F2F and VR groups ($t = 2.25$, $p = .025^{*}$, $d = .71$), but not between the F2F and VC groups ($t = .80$, $p = .42$, $d = .23$), or between the VC and VR groups ($t = 1.34$, $p = .18$, $d = .4$).

All groups' gains were negative in the post-test, and these differences proved to be statistically significant for the F2F ($t = -2.00$, $p = .042^{*}$, $d = .59$), and VR groups ($t = -2.03$, $p = .042^{*}$, $d = .64$), although Cohen's effect size values suggest moderate practical significances. On the other hand, no significant differences were found in Table 5 for the VC group ($t = -1.88$, $p = .06$, $d = .74$), which indicates higher performance for this group.

The comparison of groups' gains did not yield significant differences between groups: F2F versus VR ($t = -0.01$, $p = .99$, $d = .003$); F2F versus VC ($t = -0.2$, $p = .84$, $d = .093$); VR versus VC ($t = .2$, $p = .84$, $d = .1$).

Table 5. Accuracy: Group comparisons in presentational task

	VR		VC	
	Differences at pre-test	Differences in gains	Differences at pre-test	Differences in gains
F2F	$p = .025^{*}$	$p = .99$	$p = .42$	$p = .84$
	$d = .71$	$d = .003$	$d = .23$	$d = .093$
VC	$p = .18$	$p = .84$		
	$d = .4$	$d = .1$		

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

In summary, the three groups displayed lower performance in accuracy after the intervention, and only the VC group did not perform significantly less accurately after the treatment. This lower performance in the post-test does not necessarily mean a loss in accuracy. It could be simply a by-product of learners trying more complex structures or newly acquired contents. It is also worth noting that the standard deviations of both the VR and VC groups decreased considerably after the treatment (VC: 17.6 vs. 9.22; VR: 22.4 vs. 14). This reduction in standard deviations suggests a more even development for the voice-based CMC groups, a change not observed in the F2F group, whose standard deviation remains practically the same.

5.2.2 Interpersonal task

Group scores at pre-test ranged between 61% and 66%, but these differences were not statistically significant between any of the groups: F2F versus VR ($t = -1.51$, $p = .13$, $d = .48$); F2F versus VC ($t = 1.43$, $p = .25$, $d = .51$); VR versus VC ($t = .02$, $p = .98$, $d = .007$).

All three groups showed positive gains, but although these gains were not statistically significant for the F2F ($t = .89$, $p = .37$, $d = .18$) or VR groups ($t = 1.8$, $p = .07$, $d = .34$), they were for the VC group ($t = 3.79$, $p = .000^{\star}$, $d = .79$), with its effect size indicating a large practical significance.

The comparison between groups' gains revealed no significant differences between the F2F and VR groups ($t = -.7$, $p = .48$, $d = .19$) or the VR and VC groups ($t = 1.54$, $p = .12$, $d = .45$). The only significant difference in Table 6 was the one between the F2F and VC groups ($t = 2.25$, $p = .02^{\star}$, $d = .65$), with a moderate size effect.

Table 6. Accuracy: Group comparisons in interpersonal task

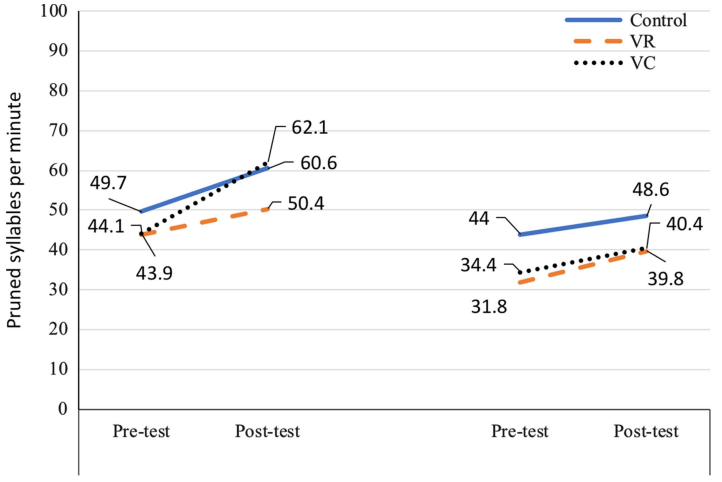
	VR		VC	
	Differences at pre-test	Differences in gains	Differences at pre-test	Differences in gains
F2F	$p = .13$	$p = .48$	$p = .15$	$p = .02^{\star}$
	$d = .48$	$d = .19$	$d = .51$	$d = .65$
VC	$p = .98$	$p = .12$		
	$d = .007$	$d = .45$		

The data indicate that completing VC activities on a weekly basis has a positive impact on accuracy on CMC interpersonal tasks, whereas doing VR and F2F activities has no effect.

5.3 Fluency

As previously noted, fluency was operationalized as the number of pruned syllables per minute. Figure 3 and the accompanying table show group mean and standard deviation values for the pre-test and post-test along with gain scores in both the presentational and the interpersonal task for all groups.

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press



	Presentational task			Interpersonal task		
	M (*SD*)	*M* (*SD*)	Gains	*M* (*SD*)	*M* (*SD*)	Gains
F2F	49.69 (16.5)	60.6 (19.3)	11.05	44 (10.9)	48.61 (9.7)	4.66
VR	43.9 (9.3)	50.37 (12.6)	6.43	31.85 (7.5)	39.78 (9.2)	7.93
VC	44.14 (12.3)	62.1 (15.2)	17.96	34.44 (12.5)	40.44 (12.3)	6

Figure 3. Descriptive statistics of fluency

5.3.1 Presentational task

Fluency scores at pre-test ranged roughly between 44 and 50, and these differences were not statistically significant between the groups: F2F versus VR ($t = -1.17$, $p = .24$, $d = .42$); F2F versus VC ($t = 1.09$, $p = .27$, $d = .37$); VR versus VC ($t = .04$, $p = .97$, $d = .019$).

In terms of changes between pre-test and post-test, the results of the analysis indicated that all groups were significantly more fluent after the intervention: F2F ($t = 5$, $p = .0000^*$, $d = .61$), VR ($t = 2.76$, $p = .0000^*$, $d = 1.29$), and VC ($t = 7.22$, $p = .0006^*$, $d = .58$), with medium effect sizes for the F2F and VC groups and a large effect size for the VR group.

The comparison between groups' gains showed in Table 7 that the VC group outperformed both the F2F ($t = 2.11$, $p = .035^*$, $d = .47$) and VR ($t = 3.38$, $p = .001^*$, $d = .77$) groups, with medium and large effect sizes respectively. No differences were found between the F2F and VR groups ($t = -1.42$, $p = .15$, $d = .31$).

Table 7. Fluency: Group comparisons in presentation task

	VR		VC	
	Differences at pre-test	Differences in gains	Differences at pre-test	Differences in gains
F2F	$p = .24$	$p = .15$	$p = .27$	$p = .035^*$
	$d = .42$	$d = .31$	$d = .37$	$d = .47$
VC	$p = .97$	$p = .001^*$		
	$d = .019$	$d = .77$		

These results suggest that regardless of treatment, beginning learners improve in fluency over a six-week period, that completing VC activities has a greater positive impact on fluency in CMC interpersonal tasks than doing VR activities or F2F activities, and that VR activities do not favor fluency development more so than F2F activities.

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

5.3.2 Interpersonal task

Mean scores at pre-test ranged between 31 and 44. The analysis revealed statistically significant differences: F2F versus VR ($t = 3.51$, $p = .000^*$, $d = 1.28$), F2F versus VC ($t = 2.66$, $p = .008^*$, $d = .82$). No differences were observed between the VC and VR groups ($t = .7$, $p = .482$, $d = .25$).

Between the pre-test and post-test, all groups' gains were statistically significant: F2F ($t = 2.76$, $p = .006^*$, $d = .44$); VR ($t = 4.48$, $p = .000^*$, $d = .94$); and VC ($t = 3.17$, $p = .002^*$, $d = .48$), with medium effect sizes for the F2F and the VC groups and a large effect size for the VR group.

The comparison of groups' gains in Table 8 showed no significant differences between groups: F2F versus VR ($t = -1.37$, $p = .17$, $d = .31$); F2F versus VC ($t = .55$, $p = .58$, $d = .13$); VC versus VR ($t = -.75$, $p = .45$, $d = .18$).

Table 8. Fluency: Group comparisons in interpersonal task

	VR		VC	
	Differences at pre-test	Differences in gains	Differences at pre-test	Differences in gains
F2F	$p = .000^*$	$p = .17$	$p = .008^*$	$p = 0.58$
	$d = 1.28$	$d = .31$	$d = .82$	$d = 0.13$
VC	$p = .482$	$p = .45$		
	$d = .25$	$d = .18$		

Overall, these results indicate that, as seen in the presentational task, beginning learners' fluency improves regardless of whether they complete F2F class activities, individual, or paired CMC voice-based activities. Furthermore, despite the fact that the F2F group had higher fluency scores before the intervention, the improvement was similar in all three groups after the six-week period.

5.4 Summary of results

Table 9 summarizes the results of the study in terms of development between the pre-test and the post-test for each group along with comparisons among groups' gains. A "+" means improvement in the post-test, "–" means a decrease in performance in the post-test, and "=" means no difference in performance in the post-test. The symbol "*" indicates a significant difference in gains between groups for each CAF condition.

Table 9. Summary of results: Within and between groups

	F2F	VR	VC
Presentational	complexity* +	complexity +	complexity* +
	accuracy –	accuracy –	accuracy =
	fluency +	fluency +	fluency* +
Interpersonal	complexity =	complexity =	complexity +
	accuracy =	accuracy =	accuracy* +
	fluency +	fluency +	fluency +

6. Discussion

This study aimed to examine the impact of two types of CMC voice-based tools on the speaking abilities of beginning learners of Spanish. The analysis of participants' performance provides insightful data in response to the research questions:

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

Research Question 1: Do VR activities help beginning learners of Spanish improve *complexity, accuracy,* and *fluency* in presentational and interpersonal modes of communication?

VR activities foster complexity and fluency in presentational tasks and fluency in interpersonal tasks. They are similar to F2F activities in both types of tasks, except regarding complexity in presentational tasks, in which F2F activities produce superior results.

These results are at odds with Guillén and Blake's findings (2017), which supported the idea that asynchronous CMC activities promote higher linguistic complexity than synchronous communication. However, the current project has important differences with that study: (1) our participants were beginning learners whereas theirs were intermediate, and (2) our study was conducted in a controlled setting and theirs was not. Namely, the participants in the Guillén and Blake study completed their activities outside of the classroom and, as a result, tended to use the so-called "best recording technique"; that is, re-recording their talk multiple times until satisfied with their production. Nation (2001) suggested this technique is a way to foster speaking fluency, but in the Guillén and Blake study, it also meant that learners had the opportunity to rehearse their performance several times, which ultimately allowed them to reach a level of complexity similar to that found in their written production. In the current study, on the other hand, performance is arguably much closer to impromptu speech, as participants were not allowed to use notes and had a limited time to form their answers.

Research Question 2: Do VC activities help beginning learners of Spanish improve *complexity*, *accuracy*, and *fluency* in presentational and interpersonal modes of communication?

VC activities promote complexity and fluency in presentational tasks and complexity, accuracy, and fluency in interpersonal tasks. When compared to the other two interactional environments, the higher levels of performance observed across the board indicate that VC activities develop different speaking performance traits than both VR and F2F activities.

These results are in conflict with those from Yanguas (2010), who found no difference between videoconferencing and F2F conversational exchanges. However, it is important to note that Yanguas' study focused on negotiation of meaning practices. Our study, on the other hand, indicates that when performance is examined in terms of CAF measures, the medium (i.e. F2F, CMC) plays, in fact, an important role in L2 speaking development. Specifically, the current study shows how participants accommodate to the demands of the medium and take advantage of its affordances. In the case of videoconferencing, there appears to be an impact on fluency in presentational tasks and on complexity and accuracy during interpersonal tasks.

Regarding fluency during presentational tasks, the results indicate that students who complete paired VC weekly activities tend to focus on the form of the message in dialogic tasks, whereas they focus more on the meaning of the message when completing monologic tasks.

The improvement in accuracy during interpersonal tasks is consistent with Xiao (2007) and can be explained in terms of the benefits of collaboration: interacting with an interlocutor promotes repair and consequently accuracy. However, the same effect was not found for the F2F group, which can be accounted for in terms of familiarity with the videoconferencing environment. Although the VC group completed VC activities every week for six weeks, the F2F group only did it twice during the whole semester. So it could be that participants in the F2F group had the additional challenge of familiarizing themselves with the medium, and as a consequence could not focus that much on the accuracy of their performance. Additionally, it is also important to take into consideration the accountability effect: it is plausible that because students in the VC group knew their performance was being recorded weekly, they trained themselves to pay attention to accuracy.

Finally, improvement in complexity during interpersonal tasks provides additional evidence of the differences between F2F and videoconferencing. In the pre-test, all groups produced numerous turns using only one or two words to acknowledge their interlocutor's responses (e.g.

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

sí, "yes"; *muy bien*, "very well"), which in videoconferencing is highly disruptive because of the characteristics of the medium (e.g. video and voice lags, delays, echo, etc.). In the post-test, only participants in the VC groups had longer AS units, characterized primarily by coordination (with *y*, "and"; *pero*, "but") and some cases of incipient subordination (*porque*, "because" to explain the previous AS unit), a change that resulted in many cases in longer turns. The fact that similar gains were not found in the F2F group suggests an accommodation to the demands of the interactional medium by beginning learners in the VC group.

Overall, the results of this study provide further empirical evidence of the connections between the use of voice-based CMC tools and the development of L2 speaking abilities, and as such can help educators make informed curricular decisions. In particular, we would like to highlight two pedagogical implications:

(1) Given the similarities between VR and F2F activities, it can be concluded that VR activities offer a suitable alternative to F2F traditional instruction for the development of CAF traits, and as such, this medium has great potential for use in online instruction. Although F2F interaction provides clear advantages over VR activities for the development of complexity in presentational tasks, it could be hypothesized that completing VR activities outside of the classroom, where students have the opportunity to rehearse as much as desired, might produce an improvement in complexity, similarly to what Guillén and Blake (2017) found.
(2) The observed advantages of VC activities over the two other mediums for the development of CAF measures in both presentational and interpersonal tasks implies that not only can they be used as an effective alternative when F2F interaction is not possible, as in online instruction, but more importantly that they should also be used as a complement to traditional instruction.

Within the larger context of SLA research, the results of this study are compatible with the main tenets of the Output Hypothesis and the Interaction Hypothesis: both individual output production and interaction with a peer of similar proficiency promote the development of language proficiency. However, the overall results also suggest that, all things being equal, interaction is a more effective means of promoting language gains.

Finally, with regard to development of CAF traits, this study provides support for Robinson's (1995) Cognition Hypothesis. In the presentational task, participants were able to focus on the meaning of the message they wanted to convey and in one of the traits of form, as shown by their improvement in both fluency and complexity. The interpersonal tasks, in contrast, foster primarily fluency, although practice in the medium of the interaction seems to have a beneficial effect in the three traits. That is, as predicted by Robinson (1995) the characteristics of the task and its administration conditions help learners overcome the trade-off effects between the different CAF components.

7. Conclusion

This study shows that medium is not merely a delivery device, but that it has important implications for learning outcomes. In this sense, the findings contribute to answering the wider question of how the use of technology in L2 instruction plays a decisive role in current teaching practices. The generalizability of the conclusions has, naturally, certain limitations. First, due to scope and space limitations, this study did not include an analysis of individual performance. Some authors (Vercellotti, 2017) have suggested that in addition to experimental conditions related to the characteristics of the task and administration conditions, it could be that some students prioritize one of the CAF traits over the others. Another limitation has to do with the demographics of the groups. Participants in the F2F group had more years of experience in

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

college than the other two groups. We acknowledge this could have impacted the results, yet in quasi-experimental studies, in which participants constitute intact class groups, researchers have no control over the demographics of such courses. Future studies could take this factor into consideration and find ways to overcome it from the onset of the study.

Ethical statement. We testify that the content of this paper presents an accurate account of the work performed as well as an objective discussion of its significance. Authors designed the materials for data collection and all participants included in this publication provided their signed consent to participate in this study.

References

Abrams, Z. I. (2003) The effect of synchronous and asynchronous CMC on oral performance in German. *The Modern Language Journal* 87(2): 157–167. https://doi.org/10.1111/1540-4781.00184

Ahmadian, M. J. & Tavakoli, M. (2011) The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency of EFL learners' oral production. *Language Teaching Research* 15(1): 35–59. https://doi.org/10.1177/1362168810383329

Beauvois, M. H. (1998) Conversations in slow motion: Computer-mediated communication in the foreign language classroom. *The Canadian Modern Language Review* 54(2): 198–217. https://doi.org/10.3138/cmlr.54.2.198

Bueno Alastuey, M. C. (2011) Perceived benefits and drawbacks of synchronous voice-based computer-mediated communication in the foreign language classroom. *Computer Assisted Language Learning* 24(5): 419–432. https://doi.org/10.1080/09588221.2011.574639

Cho, S. P. & Carey, S. (2001) Increasing Korean oral fluency using an electronic bulletin board and Wimba-based voiced chat. *The Korean Language in America* 6, 115–128.

Chun, D. M. (2006) CALL technologies for L2 reading. In Ducate L & Arnold N (eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching*. San Marcos: CALICO, 69–98. https://calico.org/product/calling-on-call/

Chun, D. M. (2011) CALL technologies for L2 reading post Web 2.0. In Arnold N & Ducate L (eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching* (2nd ed.). San Marcos: CALICO, 131–170. https://calico.org/publications/book-series/present-and-future-promises-of-call/

Elola, I. & Oskoz, A. (2011) Writing between the lines: Acquiring the presentational mode through social tools. In Arnold N & Ducate L (eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching*. San Marcos: CALICO, 171–210. https://calico.org/publications/book-series/present-and-future-promises-of-call/

Gass, S. M. & Mackey, A. (2006) Input, interaction and output: An overview. *AILA Review* 19(1): 3–17. https://doi.org/10.1075/aila.19.03gas

Gleason, J. & Suvorov, R. (2012) Learner perceptions of asynchronous oral computer-mediated communication: Proficiency and second language selves. *The Canadian Journal of Applied Linguistics* 15(1): 100–121.

Goddard, J. (1995) Perspectives on videoconferencing. *ASCILITE'95 – Learning with Technology. Twelfth Annual Conference of the Australian Society for Computers in Learning in Tertiary Education*. University of Melbourne, 3–7 December.

Guillén, G. A. & Blake, R. J. (2017) Can you repeat, please? L2 complexity, awareness, and fluency development in the hybrid "classroom. In Sanz-Sánchez I, Rivera-Mills SV & Morin R (eds.), *Online language teaching research: Pedagogical, academic and institutional issues*. Corvallis: Trysting Tree Books, 55–78.

Iwashita, N. (1999) Tasks and learners' output in nonnative-nonnative interaction. In Kanno K (ed.), *The acquisition of Japanese as a second language*. Amsterdam: John Benjamins, 31–52. https://doi.org/10.1075/lald.20.06iwa

Izumi, S. (2003) Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics* 24(2): 168–196. https://doi.org/10.1093/applin/24.2.168

Jones, L. C. (2006) Listening comprehension in multimedia environments. In Ducate L & Arnold N (eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching*. San Marcos: CALICO, 99–126. https://calico.org/product/calling-on-call/

Kern, R. (2014) Technology as *pharmakon*: The promise and perils of the Internet for foreign language education. *The Modern Language Journal* 98(1): 340–357. https://doi.org/10.1111/j.1540-4781.2014.12065.x

Long, M. H. (1996) The role of the linguistic environment in second language acquisition. In Ritchie WC & Bhatia TK (eds.), *Handbook of second language acquisition*. San Diego: Academic Press, 413–469. http://www.sciepub.com/reference/76945

Michel, M. C., Kuiken, F. & Vedder, I. (2007) The influence of complexity in monologic versus dialogic tasks in Dutch L2. *International Review of Applied Linguistics in Language Teaching* 45(3): 241–259. https://doi.org/10.1515/iral.2007.011

Murray, L. & Hourigan, T. (2006) Using micropublishing to facilitate writing in the foreign language. In Ducate L & Arnold N (eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching*. San Marcos: CALICO, 149–179. https://calico.org/product/calling-on-call/

Nation, I. S. P. (2001) *Learning vocabulary in another language*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

O'Brien, M. G. (2006) Teaching pronunciation and intonation with computer technology. In Ducate L & Arnold N (eds.), *Calling on CALL: From theory and research to new directions in foreign language teaching*. San Marcos: CALICO, 127–148. https://calico.org/product/calling-on-call/

Poza, M. I. C. (2005) *The effects of asynchronous computer voice conferencing on learners' anxiety when speaking a foreign language*. West Virginia University, unpublished PhD.

Robin, R. (2011) Listening comprehension in the age of Web 2.0. In Arnold N & Ducate L (eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching*. San Marcos: CALICO, 93–130. https://calico.org/publications/book-series/present-and-future-promises-of-call/

Robinson, P. (1995) Task complexity and second language narrative discourse. *Language Learning* 45(1): 99–140. https://doi.org/10.1111/j.1467-1770.1995.tb00964.x

Sadoux, M. (2013) *Here, there, and everywhere? The higher education academy*. http://www.hear.ac.uk/assets/documents/disciplines/Languages/Hear_there_and_everywhere.pdf.

Shehadeh, A. (1999) Non-native speakers' production of modified comprehensible output and second language learning. *Language Learning* 49(4): 627–675. https://doi.org/10.1111/0023-8333.00104

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

Skehan, P. & Foster, P. (1999) The influence of task structure and processing conditions on narrative retellings. *Language Learning* 49(1): 93–120. https://doi.org/10.1111/1467-9922.00071

Swain, M. (1985) Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In Gass SM & Madden CG (eds.), *Input in second language acquisition*. New York: Newbury House, pp. 235–256.

Swain, M. (1995) Three functions of output in second language learning. In Cook G & Seidelhofer B (eds.), *Principle and practice in applied linguistics: Studies in honour of H. G. Widdowson*. Oxford: Oxford University Press, 125–144.

Swain, M. (1997) Collaborative dialogue: Its contribution to second language learning. *Revista Canaria de Estudios Ingleses*. 34: 115–132.

Sykes, J. M. (2005) Synchronous CMC and pragmatic development: Effects of oral and written chat. *CALICO Journal* 22(3): 399–431. https://doi.org/10.1558/cj.v22i3.399-431

Vercellotti, M. L. (2017) The development of complexity, accuracy, and fluency in second language performance: A longitudinal study. *Applied Linguistics* 38(1): 90–111. https://doi.org/10.1093/applin/amv002

Wang, T. (2006) *The effects of Wimba on learning: A students and faculty perspective*. The University of British Columbia, unpublished master's thesis.

Wang, Y. (2004) Supporting synchronous distance language learning with desktop videoconferencing. *Language Learning & Technology* 8(3): 90–121.

Xiao, M. (2007) *An empirical study of using Internet-based desktop videoconferencing in an EFL setting*. Ohio University, unpublished PhD.

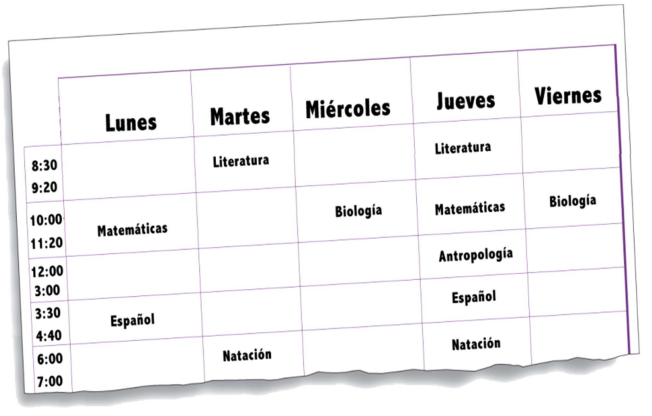
Yanguas, Í. (2010) Oral computer-mediated interaction between L2 learners: It's about time!. *Language Learning & Technology* 14(3): 72–93.

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

Appendix A

Presentational task

Your friend Verónica invited you to have lunch next Thursday. You would really like to go, but Thursday is your busiest day. You call her back to suggest another day. When you call her back, you reach her voicemail. Look at your schedule and leave her a message.



	Lunes	Martes	Miércoles	Jueves	Viernes
8:30 9:20		Literatura		Literatura	
10:00 11:20	Matemáticas		Biología	Matemáticas	Biología
12:00 3:00				Antropología	
3:30 4:40	Español			Español	
6:00 7:00		Natación		Natación	

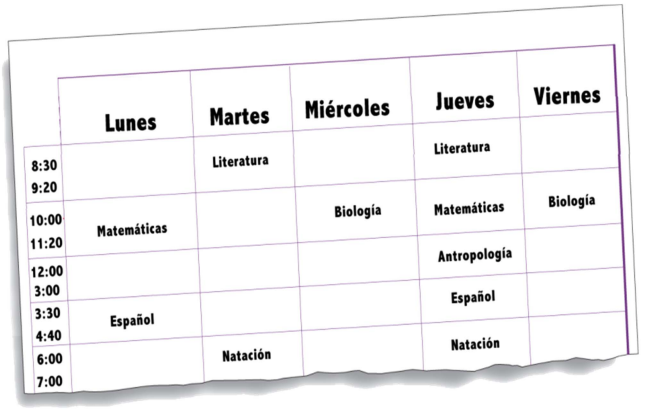
https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

- Mention how you feel about going for lunch
- Tell her about your schedule for Thursday. Include times and places
- Accept her invitation but suggest another day and time to meet for lunch

Appendix B

Interpersonal task

Estudiante A: You and your partner are good friends. It is Monday morning and (s)he calls you to invite you to go to lunch. Look at your schedule for the week below, and:



	Lunes	Martes	Miércoles	Jueves	Viernes
8:30 9:20		Literatura		Literatura	
10:00 11:20	Matemáticas		Biología	Matemáticas	Biología
12:00 3:00				Antropología	
3:30 4:40	Español			Español	
6:00 7:00		Natación		Natación	

- Greet each other
- Tell your friend about your plans for the week
- Your friend is going to invite you to have lunch this week. Accept his/her invitation but suggest another day to meet
- Suggest a specific time and a place to meet
- End the conversation properly

Estudiante B: You and your partner are good friends. It is Monday morning and you call him/her to invite him/her to go to lunch. Look at your schedule for the week below, and:

	Lunes	Martes	Miércoles	Jueves	Viernes
8:30 9:20	Tenis	Literatura	Fútbol	Literatura	
10:00 11:20	Física	Física	Física		
12:00 3:00		Historia		Historia	
3:30 4:40	Español			Español	
6:00 7:00					

- Greet each other
- Tell your friend about your plans for the week
- Invite your friend to have lunch this Thursday in a good local restaurant
- Negotiate with your partner a specific time and a place to meet
- End the conversation properly

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press

About the authors

Marta Tecedor is an Assistant Professor at Arizona State University. Her main research interest is computer-assisted language learning and instruction. In her current research she examines how the use of multimedia technologies can be incorporated in the foreign language classroom to promote the development of linguistic, interactional, and intercultural competencies.

Gonzalo Campos-Dintrans is an Assistant Professor of Spanish in the Department of Modern Languages and Literatures at the University of Mary Washington. His research focuses on adult second language learning and teaching, mainly Spanish as a second language, as well as language variation, particularly the vernacular use of the language.

Author ORCiD. Marta Tecedor, http://orcid.org/0000-0003-4546-1153
Author ORCiD. Gonzalo Campos-Dintrans, http://orcid.org/0000-0001-9016-7856

https://doi.org/10.1017/S0958344018000083 Published online by Cambridge University Press