CrossMark

ReCALL **29**(1): 59–79. 2016 © European Association for Computer Assisted Language Learning
doi:10.1017/S0958344016000161
First published online 13 September 2016

Developing oral proficiency with VoiceThread: Learners' strategic uses and views

VERA A. DUGARTSYRENOVA

National Research University Higher School of Economics, Department of Foreign Language Education, Moscow, Russia
(delvein2@gmail.com)

VERONICA G. SARDEGNA

University of Pittsburgh, Department of Instruction and Learning, USA
(vsardegna@gmail.com)

Abstract

This study explored Russian as a foreign language (RFL) learners' self-reported strategic uses of *VoiceThread* (VT)—a multimodal asynchronous computer-mediated communication tool—in order to gain insights into learner perceived effectiveness of VT for second language (L2) oral skills development and to determine the factors that contributed to those perceptions. The participants were eight undergraduate students who attended six weekly tutoring sessions that combined face-to-face (F2F) RFL instruction with VT activities. VT allowed them to access text-based, graphic, video, and aural linguistic input; record and store audio/video-recorded output; listen and comment on peers' recordings; and receive individualized teacher feedback on oral performance. Data from activity logs and researcher field notes were triangulated with participants' responses to two surveys and a semi-structured oral interview. Findings indicated that participants believed that VT supported their oral proficiency development through the provision of additional time and resources for independent planning, rehearsal, and controlled production of L2 forms. Most participants also agreed that VT's playback and record features were the most beneficial for developing language skills, enhancing the reflection process, and facilitating self-assessment and creativity in the L2. Yet, despite VT's multimodal affordances and the availability of communicative tasks via VT, participants did not perceive VT as a social environment that could promote peer-to-peer interaction or replace F2F communication. A thematic data analysis suggested that participants' preferences for language learning tasks influenced their strategic uses of VT's features, which ultimately affected their perceptions of VT's value for promoting meaningful language learning interactions. Pedagogical considerations are discussed.

Keywords: VoiceThread, Russian learners' views, language learning and teaching, computer-mediated communication, oral skills

1 Introduction

Following the surge of computer-mediated communication (CMC), the language classroom has become a fertile venue for testing out innovative technology-based projects aimed at empowering language learners and supporting their self-regulated efforts at learning

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

(Sykes, 2005; Yang, Gamble & Tang, 2011). CMC tools allow learners to plan and structure their online language learning explorations and interactions around their individual learning goals and needs. In this process, the teacher becomes a facilitator who provides language learning resources and materials, and guides students' discovery (Churcher, Downs & Tewksbury, 2014). Moreover, the affordances provided by these technologies can facilitate learners' knowledge construction and negotiation in a virtual community of practice anytime and anywhere (Bonk & Cunningham, 1998; Felix, 2005; Woo & Reeves, 2007). Yet, it is often a challenge for some educators to embed technology tools in their curricula in ways that both enhance learning and meet their students' language learning preferences. This challenge may be resolved by exploring students' self-reported uses and opinions regarding the effectiveness of technology-enhanced activities. To this aim, this study investigated how learners of Russian as a foreign language (RFL) perceived the effectiveness of one multimodal asynchronous CMC tool—*VoiceThread* (VT)—for supporting their language learning, and the factors that affected those perceptions. In this study, oral skills include the oral components of pronunciation, grammatical accuracy and syntactical complexity, vocabulary use, fluency, and clarity.

1.1 Language learning with CMC

Today's fast growing technologies offer exciting new possibilities for learning in bilingual, second, and foreign language education contexts. Some benefits of integrating CMC in these contexts include the extension of L2 oral and written discourse practices within and beyond the classroom space (Abrams, 2003; Yamada, 2009; Yaneske & Oates, 2010); the development of linguistic, intercultural, and pragmatic knowledge through instructional telecollaborative exchanges (Barron & Black, 2015; Eslami, Mirzaei & Dini, 2015); and the creation of supportive learner-centered environments that foster learners' personal engagement, self-expression, and critical and reflective thinking (Black, 2005; Hung, 2011; Sardegna & Dugartsyrenova, 2014). It has also been highlighted that the integration of CMC can remedy the "imbalances of classroom participation" (Yang *et al.*, 2011: 3) by increasing students' motivations to contribute equally online and allowing a decrease in the amount of instructor utterances (Meskill & Anthony, 2007). Scholars have also suggested that learners may feel less sensitive about saving "face" in CMC-supported environments, which could contribute to a higher degree of student involvement and self-disclosure, and a decrease in some aspects of learner anxiety (McNeil, 2014; Sykes, 2005).

Most recent research has focused on voice-based and multimodal CMC tools (e.g., virtual worlds, videoconferencing tools, voice-boards, video-blogs), which are viewed as more advantageous for language learning than just text-based CMC tools (Alameen, 2011; Hung, 2011; Lee, McLoughlin & Chan, 2008; McNeil, 2014; Nakatsukasa, 2009; Peterson, 2011; Yang *et al.*, 2011). In particular, Yaneske and Oates (2010: 241) underscore the importance of "the human element" arguing that hearing the voices of others online helped their study participants make more personal connections. Moreover, because multimodal CMC tools are freely accessible for use and build on a combination of sensory modalities (e.g., auditory and visual), they seem to be better equipped at addressing students' diverse cognitive and affective preferences and needs with regards to information processing and exchange, knowledge construction, collaboration, and social networking (Archambault & Carlson, 2011;

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Pop, Tomoletiu & David, 2011; Sun, 2009). Scholars have also posited that image and voice promote a higher level of learner consciousness of target language uses and more self-corrections (Yamada, 2009), and that multimodal tools can be effective in providing prompt teacher feedback and modeling of vocabulary use, sentence structure, and pronunciation (Yang *et al.*, 2011). Studies into virtual worlds also suggest that the combined affordances of asynchronous and real-time interaction and user-generated content offer rich opportunities for language learners to engage in novel forms of collaboration enhanced through "task-based pedagogies" (Peterson, 2011: 77), and increase students' motivations, online presence, and active participation (Deutschmann, Panichi & Molka-Danielsson, 2009; Peterson, 2011).

Some concerns regarding the value of multimodal CMC tools have also been raised. First, concerns have been linked to learners' views of voice-boards as less genuine because the oral exchanges are not produced in real time (Hung, 2011), which affects negatively the rhythm of the conversation (Yaneske & Oates, 2010). Second, studies have found that some students experience feelings of embarrassment when hearing their own voices and sharing their oral output with a larger audience (Yaneske & Oates, 2010; McNeil, 2014). Finally, voice communication has been found to have a negative effect on learners' confidence in their grammatical accuracy (Yamada, 2009).

1.2 VoiceThread and language learning

VoiceThread (www.voicethread.com) is a multimodal asynchronous CMC tool that can be used to support instruction and facilitate learning in a self-directed and safe environment. A user-generated VT presentation resembles a typical PowerPoint presentation, with each slide containing an original media artefact (a document, an image, an audio file, a video file, or a mixture of these) that serves as a prompt for discussion. Participants can offer their responses to the prompt either through file-sharing, or in the form of text, audio and/or video comments recorded and stored on the website. This variety of interactive features allows VT users to build on written, aural, and visual cues, such as facial expressions and tone of voice. Students and teachers can create their own discussions and make them accessible to the global audience or a selected number of users.

Due to its capacity for supporting computer-mediated asynchronous discussions around learners' content—uploaded or created via the medium—, VT has increasingly been used in different educational settings to support courses across a variety of disciplines, from language arts to sciences, and with diverse student populations, including learners in mainstream classrooms (Brunvand & Bird, 2011; Coiro & Kajder, 2011) and in university settings (Alameen, 2011; Kidd & Beaudry, 2013). Furthermore, scholars have argued that, in addition to fostering multimodal communicative competence and promoting new forms of social engagement (Fuchs, Hauck & Müller-Hartmann, 2012), VT increases learners' engagement with the content itself because it appeals to multiple learning modalities and allows information to be presented in digestible bits (Archambault & Carlson, 2011).

The potential of VT for language learning, and especially for promoting L2 oral and writing skills, seems promising. Yet, little is known about its pedagogical applications for language learning. The scant research that exists has mostly appeared in conference proceedings reports (Bush, 2009; Chen & Yildiz, 2010; Pallos & Pallos, 2011; Pop *et al.*, 2011;

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Sun, Yu & Gao, 2013). Findings indicate a positive reception of VT for supporting language learning and teaching. For example, teachers have used VT as a simulated conference environment to hone students' English oral presentation skills (Pallos & Pallos, 2011); as an organizational tool to support teachers when they teach poetry in their classrooms (Archambault & Carlson, 2011); as a digital story telling tool to enhance multiliteracy and collaboration skills in an advanced English as a second language (ESL) academic class (Alameen, 2011); and as a peer- and self-assessment tool for evaluating learners' L2 pronunciation, accuracy, and fluency skills development (Pop *et al.*, 2011). However, while the affordances of VT suggest it can effectively promote student learning through collaboration, self-assessment, and reflection, more empirical research is needed to validate VT's effectiveness. So far studies have reported that while VT may have increased teaching presence, it did not promote peer interaction or a sense of social community (Borup, Graham & Velasquez, 2010; Taylor & Huang, 2011). Also, Kidd and Beaudry (2013) concluded that students preferred posting text rather than voice or video comments because they found it faster, more convenient, and easier for organizing and revising their thoughts. In doing so, they denied themselves the added advantage of viewing and hearing their peers' comments. Finally, Alameen (2011: 366) found that because her ESL students relied on written scripts, their oral output on VT sounded "artificial, formal, and long" and resembled written speech. Learners were thus reported to have primarily improved on their L2 writing skills as opposed to their oral skills.

The scarce and somewhat controversial findings call for further research that could address the complexities of VT for language learning. The impact of multimodal CMC tools on students' learning gains and their oral skills is still largely unexplored, especially with regards to VT and less commonly taught world languages such as Russian.

1.3 The current study

This study attempts to add to the body of research on multimodal CMC for language learning by examining RFL students' strategic uses of VT and their opinions regarding VT's effectiveness for developing their L2 oral skills. Thus, the following research questions guided the study:

1. What are RFL learners' opinions regarding the effectiveness of *VoiceThread* for supporting their L2 oral skills development?
2. What factors influenced RFL learners' perceptions of the value of using *VoiceThread* for L2 oral skills development?

2 Method

This study employed a qualitative multiple case study approach to gain a holistic understanding of learners' experiences and perspectives (Hesse-Biber & Leavy, 2011; Stake, 1995; Yin, 2014). We chose a case study approach because it is the most appropriate to investigate how questions and a contemporary phenomenon within a real-life context (i.e., one in which researchers cannot exercise any control over the actual behavioral events) (Yin, 2014). In other words, we chose a case study approach to investigate how RFL learners used VT

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

(i.e., a real-life context) for L2 oral skills development. In keeping with a holistic approach to qualitative research, we sought to explore individual cases and describe their unique characteristics in detail, while also attempting to identify some general trends across the cases (Hesse-Biber & Leavy, 2011). To this aim, we triangulated multiple data sources.

2.1 Participants

A convenience sampling method was used to recruit participants from a first-year intensive Russian language course attended by 26 RFL learners at a large American university. To enhance students' oral skills development, additional free-of-charge technology-enhanced tutoring lessons were offered during the second semester of the course. Due to busy schedules and the fact this work did not count towards students' final course grade, only eight students from the course (seven males, one female; average age = 20) volunteered to attend the tutoring sessions. Their proficiency levels were roughly equivalent to the American Council on Teaching Foreign Languages (ACTFL) speaking proficiency scales of Novice Mid and Intermediate Low. Pre-instruction survey responses indicated that they were majoring in Linguistics (3), Economics (2), Russian Language and Culture (2), and Philosophy (1). They were all native speakers of English and none was a heritage speaker of Russian.

2.2 Context: Pedagogical design of intervention

The tutoring lessons covered three discussion themes—*College Life*, *Travelling*, and *Personalities*—and were developed and taught by a native speaker of Russian. These lessons combined online practice via VT activities with 75-minute-long face-to-face (F2F) weekly meetings with the instructor. The VT activities supported and complemented the F2F instruction, and consisted of two main types:

1. Form-focused activities targeting L2 sounds, grammatical structures, vocabulary items, and colloquial expressions.
2. Meaning-focused activities eliciting online exchanges of opinions and experiences related to the discussion themes.

Form-focused pronunciation activities were designed to enhance students' pronunciation awareness and accuracy and provide extensive practice with challenging Russian consonants (e.g., "soft" versus "hard" /л/, hissing sounds such as /ж/ "zh", /ш/ "sh", /щ/ "shch") and vowels (e.g., "soft" /и/ "i" versus its "harder" version /ы/) with no equivalent in English. These activities helped students learn to identify different sounds, and offered them opportunities for targeted oral practice through shadowing (i.e., imitating) Russian tongue twisters and rhymes. Vocabulary- and grammar-focused activities reinforced L2 forms through cloze texts and sentence completion exercises. Meaning-focused activities consisted of questions and description prompts that encouraged students to provide meaningful and clear messages as well as to exchange opinions and experiences related to the discussion themes. These prompts ranged from simple recall questions based on F2F discussions, such as "How did the characters in the story meet?" to more personalized information tasks, such as "Describe a teacher who had a major influence on you."

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

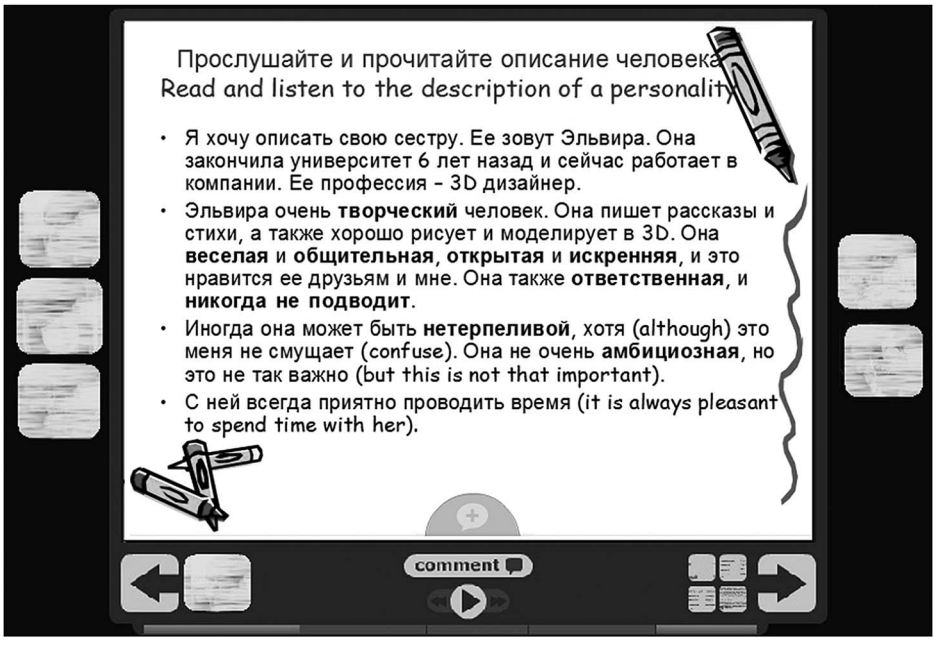


Fig. 1. Screenshot of a model text on *VoiceThread*

The multimodal online capacity of VT was also utilized to embed pre-recorded word-, sentence-, phrase-, and text-level model texts of up to 120 words into VT slides, which contained a combination of aural input, images, and textual support. Each VT presentation (total = 5) contained 10–20 slides. As participants opened the first slide of a VT presentation, they could automatically hear and replay the pre-recorded linguistic input while having its visual representation on the slide. Figure 1 displays the interface of a VT slide with a model text.

A weekly online presentation offered a variety of form-focused and meaning-focused speaking tasks. To scaffold students' learning, these tasks were presented each week in order of complexity from structured to unstructured tasks. According to Luoma's (2004: 48) classification, *structured* tasks elicit very little or no creative production (e.g., read-alouds), *semi-structured* tasks require students to produce predictable answers using aids (e.g., comprehension and sentence completion tasks), and *unstructured* tasks give learners the freedom to choose "different ways of fulfilling the task requirements" (e.g., descriptive or opinion tasks). Overall, there were 30 structured tasks (20 read-aloud and repeat-after-me tasks, and ten cloze text tasks) that focused students on form by encouraging them to reproduce pre-recorded input; 15 semi-structured tasks that provided model texts, plans, and suggested verbal structures for students to use as guides in responding to question prompts (e.g., interview a partner using question beginnings, narrate a sequence of events using sequencing words, give reasons for travel following a plan and using useful expressions) (see Figure 2); and three unstructured tasks that asked students to provide a description (see Figure 3), share a personal experience, and give instructions without any instructional aids. Both the semi-structured and unstructured tasks were mainly focused on meaning

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

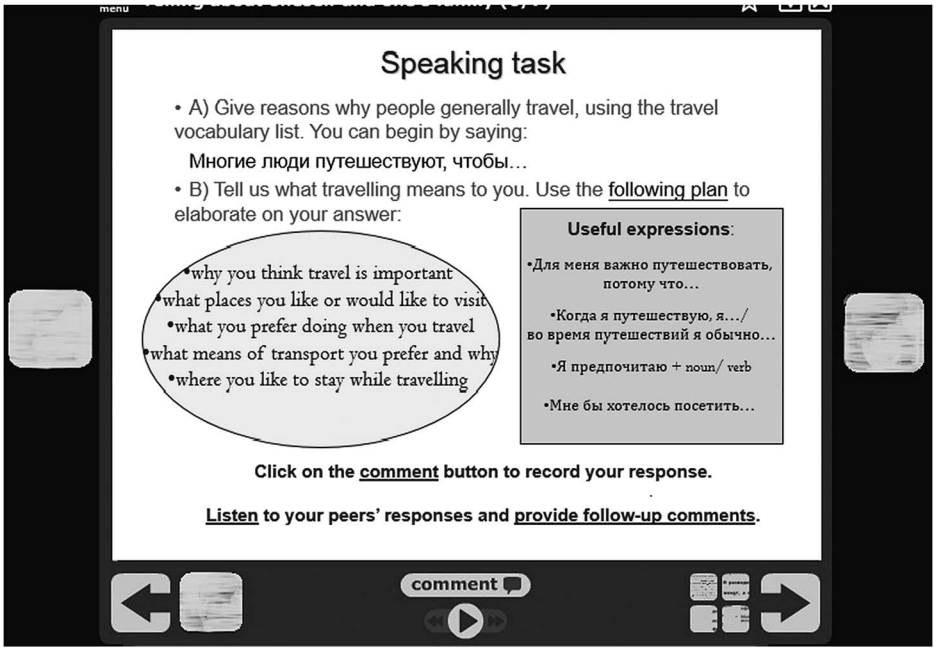


Fig. 2. Screenshot of a sample semi-structured task on *VoiceThread*

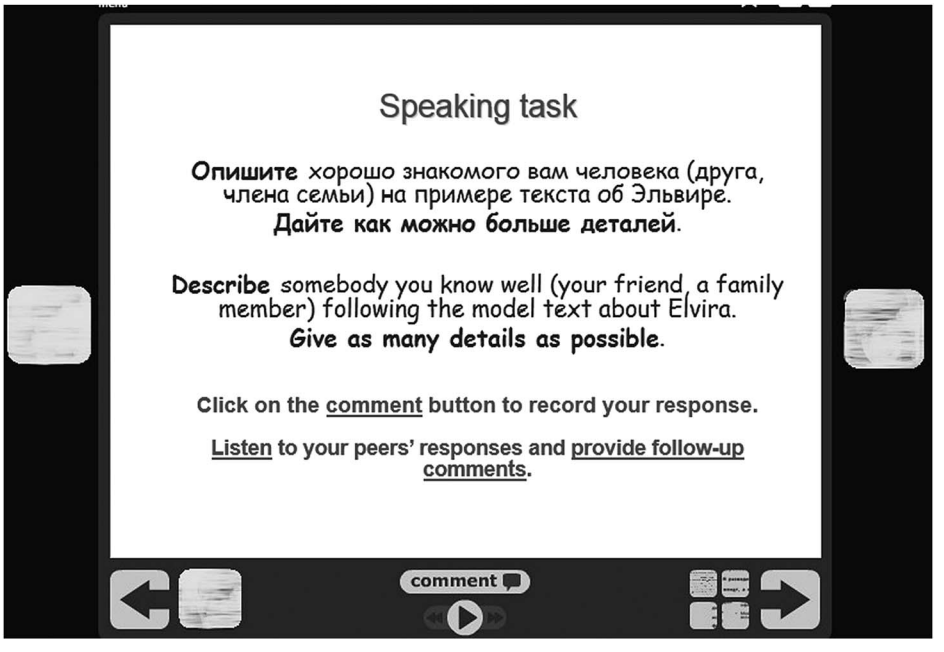


Fig. 3. Screenshot of a sample unstructured task on *VoiceThread*

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

and encouraged students to provide follow-up comments and questions on peers' posts. Participants were given the freedom to choose how and what tasks to complete each week. The only requirement was that they had to (video)record their responses and upload them for teacher feedback and peer response on the slide with the task.

Participants were also given basic instructions on how to access and use the VT platform. Captions with reminders on how to record were embedded into every slide. Students could practice using the tutor's or a peer's oral submission as a model, imitate those models or develop their own creative responses, and record themselves as many times as they wanted before publishing the final product. At the end of each week, the tutor provided individualized feedback, and commented on students' oral submissions via email and in English (i.e., students' native language). In addition to meaning-focused comments on students' posts, instructor feedback included comments on progress and suggestions for improvement in such areas as pronunciation and grammatical accuracy, word choice, and use of sentence patterns. To draw students' attention to L2 forms, the tutor provided students with grammatical explanations and/or corrections highlighted in transcribed segments from their oral messages.

2.3 Data collection and analysis

To allow data to be triangulated, multiple collection methods were used (Marshall & Rossman, 1995) (see Table 1). At week 1, all participants completed a pre-instruction online survey that elicited demographic information, language learning practice behaviors, and their opinions about own linguistic skills/challenges and best teaching and learning practices for successful language learning (see Appendix A). Evidence of participants' F2F and online engagement, practice strategies, and learning progress during the intervention was gathered through researcher field notes and an activity log of students' work. At week 6, seven of the eight participants completed a post-instruction online survey, which consisted of three open-ended questions and twelve Likert-scale items rated on a scale from 1 to 5 (from "strongly disagree" to "strongly agree") eliciting their self-reported uses of VT and their views on the benefits and limitations of VT. Finally, all students completed a focused semi-structured oral interview, which lasted approximately 20 minutes (see Appendix B). During the interview, follow-up interview questions were added when clarifications were deemed necessary. All interviews were audiotaped and transcribed for analysis.

Due to the limited sample size, inferential statistics on participants' Likert-scale survey responses were not computed. Data is better understood by analyzing the qualitative results; however, descriptive statistics on the post-instruction Likert-scale items are presented because the results (Table 2) show moderate to high agreement with all the survey items. These data were used to answer the first research question. To answer the second research question, a thematic analysis of participants' open-ended responses to the surveys and the interview was conducted using procedures adapted from a general inductive approach to coding qualitative data (Thomas, 2006). These data sources were reviewed multiple times and independently by two researchers to identify commonalities and salient themes across students' perceptions, experiences, and self-reported practices. On the basis of these generalizations, a set of categories was developed and then these categories were conceptualized into agreed-upon broader themes. Emergent hypotheses from these analyses were first tested against the data about a particular student from different data sources

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Table 1. *Data sources*

Data Source per Student	Contents	Total *N* of Sources
Tutees' recordings (Total *N* per student varied based on student choice of tasks)	Recordings of structured, semi-structured and unstructured tasks	*Aaron*: 16 (S), 14 (SS), 2 (U), T = 32* *Russell*: 10 (S), 10 (SS), 1 (U), T = 21 *Thomas*: 9 (S), 4 (SS), T = 13 *Susan*: 10 (S), 1 (SS), T = 11 *Alex*: 12 (S), T = 12 *George*: 12 (S), 3 (SS), T = 15 *Eliana*: 14 (SS), 2 (SS), T = 16 *Nick*: 10 (SS), 4 (SS), T = 14
1 Pre-instruction survey	5 background questions, 6 yes/no questions, 2 open-ended questions	8
1 Post-instruction survey	12 five-point Likert-scale items	7
1 Focused semi-structured interview	6 sample questions plus follow-up questions (duration: 20 min. approx.)	8
1 Online tracker of online activity	A log of completed VT activities per student	8
Researcher field notes (length varied per student)	Notes and comments during F2F classes Observations of online engagement	16

*S = Structured, SS = Semi-structured, U = Unstructured, T = Total tasks

and were confirmed, revised or rejected (within-case analysis). Then, these hypotheses were tested across the individual student cases (cross-case analysis) (Yin, 2014). When disagreements in data coding occurred, a third researcher was consulted. To provide a comprehensive picture, the results were triangulated with information from students' activity logs and researcher field notes.

3 Results

3.1 Perceived effectiveness of VT for oral skills development

Table 2 shows the descriptive statistics of participants' responses to the post-instruction survey. Most participants agreed or strongly agreed that while the F2F sessions increased their confidence in their oral interaction skills ($M = 4.57$), the VT presentations helped them develop their oral fluency and pronunciation skills, lexical diversity, vocabulary retention, grammatical accuracy, and syntactic complexity ($M_{range} = 4.14$–4.57). Above all, most participants appreciated their tutor's feedback on their oral recordings ($M = 4.28$). Some participants ($M = 3.86$) also indicated that they felt more comfortable posting their audio comments on VT as opposed to speaking up in class. With regards to the interaction features embedded in VT, participants noted that their tutor's and peers' online presence affected

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Table 2. *Descriptive statistics of post-instruction survey*

Statement	*Median*	*Mean**	*SD*
1. The in-class sessions with my peers and tutor helped me become more confident in communicating my ideas and asking questions in oral interaction.	5.00	4.57	1.050
2. I think that doing follow-up tasks on *VoiceThread* contributed to better retention of the vocabulary that I studied during tutoring sessions.	4.00	4.29	0.454
3. *VoiceThread* presentations helped me improve my oral fluency and pronunciation skills.	5.00	4.57	0.495
4. *VoiceThread* tasks helped me improve the lexical diversity of the sentences that I could produce in Russian.	5.00	4.43	1.050
5. *VoiceThread* tasks helped me improve the grammar accuracy and syntactic complexity of the sentences that I could produce in Russian.	5.00	4.14	1.125
6. *VoiceThread* contributed much to the improvement of my oral performance.	4.00	3.86	1.245
7. I felt more comfortable posting my audio comments on *VoiceThread* as opposed to speaking up in class.	4.00	3.86	0.990
8. I felt motivated to prepare my audio comments to *VoiceThread* tasks knowing that my oral output would be stored online and available to my instructor and my peers.	4.00	4.14	0.990
9. I found the unstructured tasks (e.g., describing a friend) the most difficult in terms of time commitment and effort to complete them.	5.00	4.00	1.195
10. The instructor's personalized comments to my performance on *VoiceThread* tasks helped me identify my oral proficiency challenges.	4.00	4.28	0.700
11. I hoped to get comments from my peers on my oral submissions.	3.00	3.00	0.535
12. Given the opportunity, I would continue doing online tasks via *VoiceThread*.	4.00	4.00	0.926

*$N = 7$

their strategy use with regard to making audio comments ($M = 4.14$); yet some wished they had received peers' comments on their recordings ($M = 3.0$). Despite the minimal student-to-student online interaction, participants generally agreed that VT contributed to improving their overall oral performance ($M = 3.86$) and that they would continue using this tool for language learning ($M = 4.0$).

3.2 Factors affecting learners' perceptions of VT's value for L2 oral skills development

Five overarching categories emerged from a cross-case thematic analysis of participants' responses to the two surveys and the interview, and researcher field notes (see Table 3): (a) preferences for tasks and practices for language learning, (b) choice of VT tasks,

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

(c) self-reported strategic uses of VT's features, (d) challenges of learning with VT, and (e) perceptions of the value of using VT for developing L2 oral skills. From the cross-case analysis, three representative cases—Aaron, Thomas, and Russell (pseudonyms)—were selected because they (a) represented other cases in terms of their preferences for language learning tasks and practices, and (b) clearly contrasted with one another in their self-reported uses of VT, which seemed to influence their challenges with VT and their views on the value of using VT for L2 oral skills development. Based on the ACTFL Proficiency Guidelines for Speaking (2012), Aaron and Russell were at an Intermediate-Low level, while Thomas was at the Novice-Mid level of Russian language competence. In order to better understand these differences, we turn to a description of these three cases. For each case and next to the participant's name, we provide the opinion statement that emerged from the within-case analysis in terms of participants' preferences for language learning tasks and practices.

3.2.1 Aaron: Languages can be learned through constant focus on form and self-correction until it becomes perfect. Researcher field notes described Aaron as extremely attentive to detail, hard-working, and very eager to learn. Aaron's responses to the pre-instruction survey indicated that he self-assessed his writing and pronunciation skills as "not bad," but reported a "quite small" vocabulary range and poor listening comprehension skills: "I need people to go slow, and then sometimes I still don't understand."

In the pre-instruction survey, Aaron reported a preference for oral activities with a potential for "real world applications" over written tasks. His activity log revealed that he attended the six tutoring sessions and chose to complete 32 VT tasks (sixteen structured, four semi-structured, and two unstructured) of the available 48 VT tasks. During the interview, Aaron admitted that he relied on two distinct learning strategies: repetition and planning. To complete a structured task, he would make as many as 20 recordings before posting his final version. He explained that he replayed a task multiple times because he "wanted to get it perfect." He described his learning process in this way:

> I'd just do the best I could and then I would, I would re-record, I would read it and then I'd try to get the intonations right, like the right part of the word intonated, but I'd try to get every word like that. (Interview)

To complete the less-structured tasks, Aaron would write down a script and consult online dictionaries to find the appropriate term although, as he himself pointed out in the interview, the online resources were not always reliable. Then, he would listen to his recording and if he did not like how it sounded, he would "re-record it, re-corded it, re-record it" until he either ran out of time or said it the way he wanted. In the post-instruction survey, Aaron stated that he tried to finish a VT presentation at one sitting, and that this task could take him approximately two hours.

When asked about the challenges he faced when using VT, Aaron replied that he struggled with the amount and complexity of the feedback he received from his tutor:

> To be honest, it brought me down sometimes. It would make me feel so bad when almost EVERYTHING I said was wrong. I worked hard on those tasks, and then I would get them back and have a million mistakes. (Post-instruction survey)

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Table 3. *Students' preferences and perceptions of VT*

	Opinion Statement	Preferences for Tasks	Choice of Tasks	Strategic Uses	Challenges	Perception of Value
Aaron	*Languages can be learned through constant focus on form and self-correction until it becomes perfect.*	Oral activities with a potential for "real world applications" and focus-on-form activities	32 VT tasks: 16 structured, 14 semi-structured, 2 unstructured	1. **Repetition**: as many as 20 versions before final posting 2. **Planning**: script preparation and use of online dictionaries before final posting	The biggest challenge was to accept the feedback from the tutor in terms of depth and quantity.	VT offers a place where one can work "constantly" and more comfortably, and be creative and resourceful.
Thomas	*Languages can be learned through listening and repetition tasks.*	Listen and repeat/imitate tasks with a focus on pronunciation accuracy	13 VT tasks: 9 structured, 4 semi-structured	1. **Repetition**: Re-recording when not satisfied with sounds, intonation, or grammatical forms produced	It was challenging to engage in a conversation and convey own ideas via VT.	VT helps improve vocabulary, grammar and listening skills; yet, above all, it helps improve pronunciation accuracy. VT is as a "place to go" to listen to native (Russian) speech.
Russell	*Language skills are best developed through face-to-face (F2F) social interactions.*	Learning the language "in real or realistic contexts" such as in classroom interactions or roleplays with peers	21 VT tasks: 10 structured, 10 semi-structured, 1 unstructured	1. **Sequencing Practice in Order of Difficulty**: From less demanding read-aloud tasks to semi-structured and unstructured tasks (in different sessions) 2. **Mental Planning**: Thinking of structures and vocabulary, and use of reference sources before posting	There is nothing limiting about VT's features. The biggest challenge is not to lose focus while working on VT tasks.	VT offers the opportunity for practice, "seeing things and hearing them at the same time," and targeted tutor feedback; but it does not offer engagement with social activities typical of F2F interactions.

Aaron further admitted that he found his tutor's feedback "hard to endure" during the first tutoring weeks; yet, as the weeks progressed, he felt much better about it: "I just kept plugging away though, and the mistakes got less by the end of the year. Now, the feedback isn't that bad at all." In the interview, Aaron reported also struggling with VT's intrinsic lack of capacity to provide real-time or "live" feedback on his performance.

When asked during the interview about the value of using VT for L2 oral skills development, Aaron indicated that VT helped him "put his classroom knowledge into real use." He found VT assignments especially rewarding because they enabled him to discuss issues that went beyond class topics and were personally relevant. Furthermore, although he noted that the unstructured tasks were time-consuming, Aaron felt that they pushed him to be creative and resourceful: "With VT, it was like the dam of the Russian language broke and flooded over me. There were no more little pieces to learn. Everything was available to me" (interview). Aaron also stated that the F2F sessions were more rewarding in terms of helping him compare his progress to that of his classroom peers. Yet, for him, VT was a place where he could work "constantly" and more comfortably. In other words, Aaron seemed to appreciate VT's asynchronous nature because it allowed him to repeat and plan his oral submissions at his own pace and as much as he needed. The following quote from the post-instruction survey summarizes Aaron's overall perception of his progress at the end of the project: "Now I can honestly say I feel I am one of the best students in the class, and I can attribute at least 75% of that to VT and X [the tutor] for making me not stay complacent."

3.2.2 Thomas: Languages can be learned through listening and repetition tasks. In the pre-instruction survey, Thomas indicated that he believed his strongest skill was pronunciation. He praised his ability to memorize vocabulary, while noting that his speaking and grammar were "awful." He found it a challenge to "have a simple conversation" and convey his ideas orally or in writing. Furthermore, Thomas exhibited a preference for engaging in listen and repeat/imitate tasks because, as he explained during the interview, this type of task is not only fun and less challenging but also "helps with your pronunciation." Thomas stated in the pre-instruction survey that he needed "several sessions" (i.e., repetition) to "feel comfortable using a concept while speaking." When asked during the interview about the different VT tasks, Thomas admitted:

> To be honest, the ones that would probably give me the most are the ones when you have to think of original answers … Yeah, although I don't really enjoy those because it's hard. I liked the tongue twisters—that kind of helps you with speaking and it's kinda fun trying to see how quickly you can go through them all being accurate. Those are good. (Interview)

Based on his activity log, Thomas attended four of the six tutoring sessions, and chose to complete only thirteen VT tasks (nine structured and four semi-structured). During the interview, Thomas explained that he completed so few because he had other commitments that required his attention.

In terms of his learning strategies, Thomas reported that he would re-record tasks when he did not like the way he articulated certain sounds, was not content with his intonation, or was aware of his grammatical mistakes. In the post-instruction survey, Thomas admitted

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

that on average he would spend an hour a week on VT tasks, but if he had completed every task, "it probably would've been at least two hours or so a week."

Furthermore, Thomas indicated that he focused on improving his pronunciation accuracy because he cared about what his audience would think or feel when they listened to his oral speech:

> I'd definitely rehearse … 'cause I wanted it to … I mean I wanted to say it right just so I could give my best effort, and I don't know… just so it wouldn't be a bunch of pauses and me correcting myself. That'd be, kinda, sound annoying … having to listen to that. (Interview)

Thomas's awareness of his audience also became evident when he admitted during the interview that he took the time to replay his tutor's and peers' posts to use them as points of reference, even though he found his peers' posts "not nearly as beneficial" as his tutor's posts.

When asked about the challenges he faced when using VT, Thomas explained that he favored the tutoring sessions over VT because he saw them as more efficient: "They [the F2F lessons] do a lot for me. Unfortunately, I was unable to put into these sessions the time and attention that would have made them more effective" (interview). Nonetheless, he still referred to VT as a priority "place to go where [he] could listen to different Russian phrases said aloud" (interview). Thomas also seemed to appreciate and welcome any kind of feedback, as indicated by: "I invite all types of critiques in response to my recordings."

When asked during the interview about the value of using VT for L2 oral skills development, Thomas reported that he felt VT had helped him improve his listening skills and use of vocabulary and grammatical structures. In the post-instruction survey, he also pointed out that his major achievement on VT as compared to that of his online peers was an increase in his pronunciation accuracy: "The one thing that I feel that I am better at than most people are is pronunciation."

3.2.3 Russell: Language skills are best developed through face-to-face social interactions. Researcher field notes described Russell as the most proficient of the group. He already spoke four languages with varying degrees of fluency. His responses to the pre-instruction survey indicated that he believed his "passive skills" were better developed than his "active skills" and that his "graphic skills" (reading comprehension) were better than his "acoustic skills" (listening comprehension). Although he thought he had a relatively large vocabulary, Russell did not see himself as a fluent speaker in terms of producing "spontaneous speech."

In the pre-instruction survey, Russel indicated preferences for "fully analyzing native materials such as songs and news articles with related vocabulary and grammar;" yet he thought that learning the language "in real or realistic contexts" was most beneficial for his language development. Furthermore, his interview responses revealed that he preferred the F2F tutoring sessions to the practice on VT for enhancing his speaking skills. For Russell, effective conversation practice occurred in settings such as "classroom communication and instruction, simulated situations acted out with peers, and short dialogues in pairs" (pre-instruction survey). During his interview, Russell explained that he believed that these settings were more efficient for developing oral proficiency because of "the faster pace of

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

understanding and responding necessary to sustain a lively conversation" and because "there is more pressure to produce language quickly face-to-face and more immediate feedback on pronunciation and grammatical accuracy." Yet, as Russell also admitted in his interview, it was challenging for him to engage in conversation practice, and he did not view VT as a platform that would support such practice:

> It's a bigger challenge for me to carry on a conversation with somebody in the language, something I guess I need to work on, but I don't know, so it [VT] was really, it was not something that I really viewed in a social way I guess. (Interview)

Russell's activity log indicated that he attended all the tutoring sessions and chose to complete 21 VT tasks (ten structured, ten semi-structured, and one unstructured). When the interviewer asked him how he approached tasks on VT, Russell described his learning strategies as a series of steps completed during several sessions. He would first start with the less demanding read-aloud tasks, reserving the semi-structured and unstructured tasks for a different session. These, as he said, "generally take more time than just reading through the words that are already there." In approaching the more demanding tasks, Russell engaged in mental planning of what he intended to say. Before hitting the comment (record) button, he would think about specific structures and vocabulary to use in his response, and would resort to reference sources (e.g., dictionaries) if needed:

> I would generally just think through it mentally and kind of try to plan out what I was going to say and made sure I knew all the structures that I wanted to use, and then if there's any particular vocab that I wanted to use then I'd go look it up in a dictionary, and then I would just record it. (Interview)

After submitting his first recording, Russell would replay it to see what revisions were needed: "Generally the first time would be not very good, I'd make some mistakes or stumble, but then I'd re-record it once more, and it could generally be fluent and more sound good" (interview). In the post-instruction survey, Russell indicated that his submissions took him five to ten minutes per task.

When asked to identify challenges, Russell observed, "Nothing particularly stood out to me. I don't know. I don't think there was anything that was particularly limiting about the format of VT" (interview). Yet, his comments suggested a preference for learning in a classroom-based environment, as shown in the following excerpt:

> Mostly I find it easier to stay on task when I am in the class, doing something, because a lot of times when I am on the computer I have my email up, or Facebook, or music. There's just lots of things distracting me, so I end up getting through the VT assignments pretty slowly. (Interview)

In terms of the value of using VT for L2 oral skills development, Russell indicated in the post-instruction survey that he appreciated the opportunity for "just any speaking practice, in any context" afforded by VT tasks. He also valued "seeing things and hearing them at the same time" and receiving feedback from his tutor. Reflecting on his oral skills, he noted an overall improvement on fluency, vocabulary and grammar, and a stronger confidence in his ability to speak Russian both in and outside the classroom. During the interview, he stated,

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

> I liked the opportunity to mentally compose a long response and then record it, listen to it and re-record it to correct my errors and make it sound more fluent. I feel like this provided a stronger reinforcement for grammatical structures than extemporaneous speech in which it is not possible to go back and correct one's errors. (Interview)

However, when gauging the overall impact of VT on his progress in the larger context of his other language activities, Russell described the online component as:

> A small but significant portion of the work I did on Russian throughout the year. Most of my progress in the language has come from exposure to language outside of class, as I often listen to music, surf the web, and watch videos and movies in Russian. (Post-instruction survey)

When asked to indicate in the post-instruction survey whether he would continue using VT for oral proficiency development and whether he would recommend it to a friend, Russell chose to state that he would do neither. Thus, it seems that his strong preference for constant F2F exposure to the L2 and for engaging in social activities for language learning overrode his perceptions of language gains via VT.

4 Discussion

The findings revealed that participants were positive about the affordances of VT for developing their oral skills, and most would consider using VT again for language learning. In particular, they valued the opportunity to listen to native speech and imitate it, the extensive planning and delivery time afforded by VT's embedded features, and the chance to self-assess and re-record themselves as needed to refine their oral submissions. The results also indicated that such focused practice, and learners' ongoing reflection and self-monitoring increased most participants' confidence in different areas of oral performance—i.e., oral fluency and pronunciation skills, lexical diversity, vocabulary retention, grammatical accuracy, and syntactic complexity—as well as their comfort with sharing their oral output with others. Some even commented that they felt more comfortable posting their audio comments on VT than speaking up in class. Additionally, participants appreciated the scaffolding provided by the tutor through recorded input, targeted and delayed feedback, and F2F instruction. Even Aaron, who initially struggled due to the amount of errors he needed to correct based on tutor feedback, later admitted that such feedback motivated him to work harder.

Despite participants' positive reception of VT, many underscored the need to combine out-of-class VT practice with F2F tutoring. These participants felt that the F2F instruction was necessary because it increased their opportunities for immediate feedback and oral fluency development. While a few acknowledged that they could listen to their peers' recordings and many said that they would have liked to receive comments from peers via VT, only a handful (e.g., Thomas) took advantage of VT's multimodal features and listened to peers' posts. Interestingly, even when many of the tasks encouraged them to post comments and interact with other students online, none chose to comment or provide follow-up feedback. Not even participants like Russell, who voiced a preference for learning through "talking" with others, took advantage of VT's interactive features. A possible explanation for this outcome is that VT's interactive features did not afford "real-time" interaction or

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

immediate feedback on oral production, which many participants considered essential for language learning. Also, based on students' comments, immediate feedback seemed to have been effectively provided in the F2F tutoring sessions, which may have ultimately contributed to a decrease in students' perceived need for online peer interaction and/or feedback, and a preference for a focus on forms. Unfortunately, by choosing to interact with peers only during F2F sessions, participants denied themselves the benefits of peer-to-peer online collaboration, sharing, and feedback (also found in Borup *et al.*, 2010).

Finally, our cross-case analysis of participants' self-reported strategic uses of VT revealed that their preferences for language learning activities manifested themselves in their choices of VT tasks and their strategic uses of VT features. For example, Aaron's preference for accuracy tasks and Thomas's preference for listen and repeat tasks influenced their decisions to complete most of the structured tasks, and to engage in multiple recordings following speech models—i.e. they resorted to using repetition and imitation strategies to improve their oral skills. In contrast, Russell's preference for learning through using the language in interaction manifested itself in his choice for completing the structured tasks with a focus on fluency and clarity. That is, he would only rerecord his output if he felt it was not fluent or clear enough. Also, because Aaron was more focused on accuracy, he developed a script for the semi-structured and unstructured tasks, which he also rerecorded many times before submitting it. In contrast, Russell would think what to say first (i.e., focus on meaning) before recording his work, and Thomas skipped the unstructured tasks altogether to have more time for listen and repeat tasks. In turn, their choices determined the challenges they faced while learning with VT. For example, Aaron, who focused on accuracy, had a much harder time accepting tutor feedback than Russell, who focused mostly on communicating a clear and fluent message. Ultimately, these participants' task choices, self-reported uses of VT's features, and challenges affected their overall perceptions of the value of using VT for developing their L2 oral skills. Overall, the results revealed that learners with a preference for independent practice, who also chose to do the tasks that required such practice (like Aaron and Thomas), highly valued the use of VT for oral skills development. However, learners with a preference for learning through interaction, but who did not use VT in that way (like Russell), ultimately did not think VT was an effective tool for promoting their oral skills. Thus, it seems of vital importance that teachers consider the effects of students' preferences and learning styles when using multimodal tools in self-directed learning. Because the interaction tasks were the most challenging and were not compulsory, they were also the least frequently chosen, which left students like Russell, who wanted to engage in interactions, with little to no opportunities for finding a conversation partner.

Our findings suggest some pedagogical considerations that may be helpful for teachers interested in enhancing learning L2 oral skills with VT. First, explicit technical guidelines for working with VT's main features should be offered throughout the intervention to ensure that the potential of VT's affordances is maximized. This guidance could include information regarding the benefits of using communicative tasks and VT's multimodal features (e.g., the webcam) for language learning. It is possible that drawing students' attention to these tasks and features might increase students' uses of them. Second, teachers should be aware that technical knowledge and opportunities to use multimodal tools do not necessarily result in students' making use of these tools to interact with each other, or to co-construct knowledge within a community of practice (Bonk & Cunningham, 1998).

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

As our research and other studies (Borup *et al*., 2010; Taylor & Huang, 2011) suggest, VT per se does not promote peer interaction. If teachers require, model, and scaffold peer-assessments with guidelines for constructive feedback, they may minimize negative feelings (e.g., losing confidence) and lower aspects of learner anxiety that have been found to hinder students' interactions in voiced-CMC tasks (see McNeil, 2014; Sykes, 2005; Yamada, 2009). Third, to scaffold students' efforts at learning, teachers should provide sufficient linguistic support, explicit instruction, and targeted feedback. As revealed by our analysis, participants appreciated and were motivated by the ongoing tutor feedback. However, for the feedback to be most effective, it should be provided in manageable bits and according to students' needs and abilities so that it does not overwhelm or stress them unnecessarily, as it occurred with Aaron. Finally, if combined with F2F interactions or synchronous CMC exchanges, VT's potential for collaborative language learning might be enhanced. F2F and synchronous online communication could compensate for students' perceived lack of spontaneous interaction and feedback on VT and VT's assumingly "weak linkage to real-time communication" (Hung, 2011: 744).

4.1 Limitations and future research

Due to its exploratory nature, this study presents some limitations. It is based on the opinions of a small student population that chose to take the tutoring classes. It is possible that their motivations for out-of-class practice and interest in learning with technology may have increased the likelihood of a positive response to VT. In addition, these students volunteered to participate in extra activities so they had more opportunities to learn Russian than the rest of the class. It is also highly probable that they were also more motivated to learn Russian than the other members of the class. Future studies might include populations without such inclinations. Also, to assess participants' actual language gains, future research could analyze their oral performance through a pre-/post-test design. With regards to the potential of VT for promoting tutee-tutor and peer-to-peer interaction, future studies might want to investigate whether required and modeled peer-feedback would increase students' social presence online. Finally, an investigation of the tutor's strategies for providing feedback directly via VT may shed light on its effect on learning.

5 Conclusion

This study offered preliminary insights into the pedagogical applications of a multimodal CMC tool—*VoiceThread*—for supporting oral proficiency development for learners of RFL. During six weeks, participants were instructed to complete form-focused and meaning-focused structured, semi-structured, and unstructured tasks embedded in VT presentations for individualized oral practice, and to use any of the tool's features (e.g., playback, recording, file-sharing) to support their development of L2 oral skills. The findings suggested that participants most highly valued VT's affordances for repetitive playback, read-aloud rehearsal, planning, reflection on progress, and comparison of own oral output against that of the tutor. These affordances allowed them to identify tasks and areas for focused practice, enhance their self-assessment and awareness of own progress, engage in more creative L2 uses and, ultimately, increase their confidence in their oral skills.

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Yet, while most learners appreciated the benefits of using VT for improving their pronunciation, grammar, and vocabulary, they did not regard VT as a space that encouraged peer-to-peer interaction and could effectively support conversational exchanges in L2, as compared to traditional classroom venues. A thematic analysis of the findings revealed that participants' self-reported strategic uses of VT's features were influenced by their preferences for language learning tasks and practices, which ultimately affected their perception of the value of using VT for L2 oral skills development. Pedagogical considerations were drawn for maximizing VT's multimodal affordances, enhancing the development of L2 oral skills, and enriching learners' experiences using VT.

References

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

Alameen, G. (2011) Learner digital stories in a Web 2.0 age. *TESOL Journal*, **2**: 355–369.

American Council on the Teaching of Foreign Languages (2012) *The ACTFL proficiency guidelines for speaking*. http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/english/speaking

Archambault, L. and Carlson, D. (2011) Poetry in motion: using VoiceThread to prepare 21st century English teachers. In: Koehler, M. and Mishra, P. (eds.), *Society for Information Technology & Teacher Education International Conference*. Nashville, Tennessee, USA: Association for the Advancement of Computing in Education (AACE).

Barron, A. and Black, E. (2015) Constructing small talk in learner-native speaker voice-based telecollaboration: a focus on topic management and backchanneling. *System*, **48**: 112–128.

Black, A. (2005) The use of asynchronous discussion: creating a text of talk. *Contemporary Issues in Technology and Teacher Education*, **5**(1): 5–24.

Bonk, C. J. and Cunningham, D. J. (1998) Searching for learner-centered, constructivist, and sociocultural components of collaborative educational tools. In: Bonk, C. J. and King, K. S. (eds.), *Electronic collaborators: learner-centered technologies for literacy, apprenticeship, & discourse*. Mahwah, NJ: Erlbaum.

Borup, J., Graham, C. and Velasquez, A. (2010) The use of asynchronous video communication to improve instructor immediacy and social presence in an online course. In: Gibson, D. and Dodge, B. (eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2010*. Chesapeake, VA: AACE.

Brunvand, S. and Bird, S. (2011) Using VoiceThread to promote learning engagement and success for all students. *Teaching Exceptional Children*, **43**(4): 28–37.

Bush, L. (2009) Viva VoiceThread: integrating a web 2.0 tool in the additional language classroom. In: Crawford, C., Carlsen, R., Gibson, I., McFerrin, K. and Willis, D. (eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2009*. Chesapeake, VA: AACE.

Chen, J. and Yildiz, M. (2010) Preparing English language learners for academic success in the 21st century: teaching multiple literacies. In: D. Gibson and B. Dodge (eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference*. Chesapeake, VA: AACE.

Churcher, K. M. A., Downs, E. and Tewksbury, D. (2014) "Friending" Vygotsky: a social constructivist pedagogy of knowledge building through classroom social media use. *The Journal of Effective Teaching*, **14**(1): 33–50.

Coiro, J. and Kajder, S. (2011) Conversation currents: digital footprints. Published interview and podcast. *Language Arts*, **89**(2): 148–153.

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Deutschmann, M., Panichi, L. and Molka-Danielsson, J. (2009) Designing oral participation in Second Life: a comparative study of two language proficiency courses. *ReCALL*, **21**: 206–226.

Eslami, Z., Mirzaei, A. and Dini, S. (2015) The role of asynchronous computer mediated communication in the instruction and development of EFL learners' pragmatic competence. *System*, **48**: 99–111.

Felix, U. (2005) E-learning pedagogy in the third millennium: the need for combining social and cognitive constructivist approaches. *ReCALL*, **17**(1): 85–100.

Fuchs, C., Hauck, M. and Müller-Hartmann, A. (2012) Promoting learner autonomy through multiliteracy skills development in cross-institutional exchanges. *Language Learning & Technology*, **16**(3): 82–102.

Hesse-Biber, S. N. and Leavy, P. (2011) *The practice of qualitative research* (2nd edn.). Los Angeles: Sage.

Hung, S.-T. (2011) Pedagogical applications of Vlogs: an investigation into ESP learners' perceptions. *British Journal of Educational Technology*, **42**(5): 736–746.

Kidd, J. and Beaudry, J. (2013) Understanding students' online communication preferences and the affordances of VoiceThread for formative assessment in online teaching. In: McBride, R. and Searson, M. (eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2013*. Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Lee, M., McLoughlin, C. and Chan, A. (2008) Talk the talk: learner-generated podcasts as catalysts for knowledge creation. *British Journal of Educational Technology*, **39**: 501–521.

Luoma, S. (2004) *Assessing speaking*. Cambridge: Cambridge University Press.

Marshall, C. and Rossman, G. B. (1995) *Designing qualitative research* (2nd edn.). Thousand Oaks, CA: Sage.

McNeil, L. (2014) Ecological affordance and anxiety in an oral asynchronous computer-mediated environment. *Language Learning & Technology*, **18**(1): 142–159.

Meskill, C. and Anthony, N. (2007) Form-focused communicative practice via CMC: what language learners say. *CALICO Journal*, **25**(1): 69–90.

Nakatsukasa, K. (2009) The efficacy and students' perceptions of collaborative blogging in an ESL classroom. In: Chapelle, C. A., Jun, H. G. and Katz, I. (eds.), *Developing and evaluating language learning materials*. Ames, IA: Iowa State University.

Pallos, H. and Pallos, L. (2011) Evaluation of Voicethread technology to improve Japanese graduate students' presentation skills in English in a blended learning environment. In: Barton, S., Hedberg, J. and Suzuki, K. (eds.), *Proceedings of Global Learn Asia Pacific 2011*. AACE.

Peterson, M. (2011) Towards a research agenda for the use of three-dimensional virtual worlds in language learning. *CALICO Journal*, **29**(1): 67–80.

Pop, A., Tomoletiu, E. and David, D. (2011) EFL speaking communication with asynchronous voice tools for adult students. *Procedia – Social and Behavioral Sciences*, **15**: 1190–1203.

Sardegna, V. G. and Dugartsyrenova, V. A. (2014) Pre-service foreign language teachers' perspectives on learning with technology. *Foreign Language Annals*, **47**: 147–167.

Sun, Y. C. (2009) Voice blog: an exploratory study of language learning. *Language Learning & Technology*, **13**(2): 88–103.

Sun, Y., Yu, J. and Gao, F. (2013) Shared video media: A new environment to support peer feedback in second language learning. In: McBride, R. and Searson, M. (eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2013*. Chesapeake, VA: AACE.

Stake, R. (1995) *The art of case study research*. Thousand Oaks, CA: Sage.

Sykes, J. (2005) Synchronous CMC and pragmatic development: effects of oral and written chat. *CALICO Journal*, **22**(3): 399–431.

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press

Taylor, L. and Huang, H. W. (2011) Student engagement in online multimedia communication. In: Koehler, M. and Mishra, P. (eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2011*. Chesapeake, VA: AACE.

Thomas, D. R. (2006) A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, **27**: 237–246.

Woo, Y. and Reeves, T. C. (2007) Meaningful interaction in web-based learning: a social constructivist interpretation. *The Internet and Higher Education*, **10**: 15–25.

Yamada, M. (2009) The role of social presence in learner-centered communicative language learning using synchronous computer-mediated communication: experimental study. *Computers & Education*, **52**(4): 820–833.

Yaneske, E. and Oates, B. (2010) Using Voice Boards: pedagogical design, technological implementation, evaluation and reflections. *Research in Learning Technology*, **18**(3): 233–250.

Yang, Y.-T., Gamble, J. and Tang, S.-Y. (2011) Voice over instant messaging as a tool for enhancing the oral proficiency and motivation of English-as-a-foreign-language learners. *British Journal of Educational Technology*, **43**(3): 448–464.

Yin, R. K. (2014) *Case study research: design and methods*. Thousands Oaks, CA: Sage Publications.

Appendices

Appendix A: Pre-instruction online survey

Please provide the following information about you.
Gender: Age: Nationality: Major(s):

- When did you start learning Russian?
- What do you do to improve your Russian language skills?

Watch Russian films	Yes/No
Listen to Russian performers	Yes/No
Read books in the Russian language	Yes/No
Browse Russian websites	Yes/No
Meet with Russian-speaking friends	Yes/No
Chat online with native speakers	Yes/No

- What language teaching practices are important for you to be a successful language learner?
- Which Russian language skills (reading, writing, listening, speaking, grammar, vocabulary, pronunciation) do you think are your strongest? Which do you need to improve?

Appendix B: Sample oral interview questions

- How does working with *VoiceThread* compare to a traditional foreign language class?
- What kind of tasks on *VoiceThread* were the most/least effective? Please provide examples and explain why.
- What steps did you take to record your verbal responses to the different tasks? How much time, on average, did you spend on recording such responses?
- Which features of *VoiceThread,* if any, have you found especially beneficial for promoting oral language development?
- What were the major challenges of working with *VoiceThread*?
- What are *VoiceThread*'s major limitations in terms of fostering oral proficiency?

https://doi.org/10.1017/S0958344016000161 Published online by Cambridge University Press