TEACHING with TECHNOLOGY 2020

Selected papers from the JALTCALL2020 Conference

JALTCALL 2020

Edo Forsythe, Editor-in-Chief

Associate Editors: Anthony Brian Gallagher
Douglas Jarrell
Branden Kirchmeyer
Louise Ohashi
James York

ISBN 978-4-901352-63-5

Teaching with Technology 2020
Selected papers from the JALTCALL2020 Conference

JALT CALL is a Special Interest Group (SIG) of the Japan Association for Language Teaching (JALT) that focuses on Computer-Assisted Language Learning (CALL) and technology in language learning.

http://jaltcall.org/

JALT is dedicated to the improvement of language teaching and learning.

JALT Central Office
Urban Edge Bldg. 5F, 1-37-9 Taito,
Taito-ku, Tokyo 110-0016, Japan

http://jalt.org/

First published in Japan in 2021 by the JALT CALL SIG.

Layout by Paul Mason

Preface

JALTCALL2020 broke barriers in a number of ways, but one of the key points to this year's conference was a back-to-basics focus on using technology in language teaching and learning. The online conference allowed participants and presenters from around the world to share their ideas and experiences with one another. This year's topics ranged widely with presentations and posters giving tips for teaching all four language skills and a multitude of content topics using a variety of platforms and apps. The conference was highlighted by Dr. Charles Browne's Keynote presentation in which he introduced the New General Service List (NGSL) Project and explained how it can be used to help learners focus their vocabulary learning efforts. Because the global COVID-19 pandemic forced many language teachers to teach online, JALTCALL2020 proved to be an outstanding trove of resources to make the transition to teaching with technology easier and more productive. This volume of collected papers provides a good sample of the research and practice pertaining to the conference theme. An invited chapter is included based on Dr. Browne's keynote presentation. Also, eight papers were chosen to be included in this work and they feature wide-ranging topics such as taking an English curriculum online, teaching writing using mobile devices, the use of virtual reality systems in language learning, student collaboration using digital backchannels, teaching law students to write better in English, and a survey of current MALL research in Japan. The papers contained herein have been double-blind peer reviewed and chosen for publication according to their quality, suitability for inclusion, and academic relevance to CALL research.

The editors worked with the authors to improve their manuscripts for publication in this volume of collected papers. I hope that the readers of this volume will find these papers insightful, useful, and practical for language teaching in Japan and around the world.

I would like to thank the authors who worked diligently to provide well-researched studies and practically-focused papers with the common theme of CALL. Additionally, this book would not be possible without the dedication of our volunteer associate editing staff. I greatly appreciate their efforts in working with the authors to hone their manuscripts and prepare this volume for publication.

Editor-in-Chief
Edo Forsythe, EdD, Hirosaki Gakuin University

From the JALTCALL2020 Co-Chairs

At the start of 2020, we had no inkling that JALTCALL2020 would be the first online conference for the SIG. However, through January and February global events came one one after the other: bush fires, assassinations, and the beginnings of a global pandemic. As we headed into spring, the possibility of holding a face-to-face conference dwindled. The CALL SIG officers came together to plan an alternative. That alternative was to host an online conference using Zoom and a custom conference tool designed by Gary Ross, co-chair of JALTCALL2020. Not only was the online conference the first of its kind for the SIG and for JALT at large, it was also the first conference which was free to attend for attendees, JALTCALL2020 essentially served as the template foundation for the following conferences: PanSIG 2020, the 11th Annual Extensive Reading in Japan Seminar, and the JALT2020 International Conference. This allowed for the widest possible dissemination of ideas, and fantastic networking opportunities as we had over 600 attendees from about 50 countries – such as Singapore, Turkey, Germany, Myanmar, Finland, Ireland, Argentina, Algeria, Nepal, and Reunion – join us for the conference.

This book is a collection of papers based on presentations from JALTCALL2020. We hope attendees had a positive experience at the conference and are incredibly grateful to 1) all SIG officers for pulling together and creating a fabulously successful conference amid global turmoil, and 2) everyone that presented and attended. We hope to see you at JALTCALL2021 on June 4–6, 2021 at Hirosaki Gakuin University in beautiful western Aomori.

James York, Tokyo Denki University
Gary Ross, Kanazawa University

Editorial Team

Editor-in-Chief

Edo Forsythe is a professor and head of the Department of English Language and Literature at Hirosaki Gakuin University. His research interests include MALL and the use of smartphones in language learning. Twitter: @Prof_Edo

Associate Editors

Anthony Brian Gallagher is a lecturer in the Faculty of Foreign Studies at Meijo University. His research interests are quality assurance of the user experience; assistive, disruptive, & cosmetic technologies; and borderline learning disabilities.

Douglas Jarrell is a retired professor, formerly at Nagoya Women's University, whose research interests include the affordances of technology for self-study and student engagement in the classroom.

Branden Kirchmeyer is a lecturer at Sojo University in Kumamoto. He is currently researching and developing strategies for formative assessment and learner profiling for foreign language learning.

Louise Ohashi is an associate professor at Meiji University who specialises in English language education. Her doctoral work and much of her other research examines teacher support of learner autonomy, learner motivation, and the use of digital technologies for language learning. Twitter: @OhashiLou

James York is a lecturer at Tokyo Denki University where he researches the application of games and play in language teaching contexts. He also co-edits *Ludic Language Pedagogy,* an open-access journal. Twitter: @cheapshot

Layout

Paul Mason, Aichi Gakuin University

Technical Editor (website programming and design)

Gary Ross, JALT CALL SIG Webmaster

Contents

1 The NGSL Project: Building Wordlists and Resources to help EFL Learners (and Teachers) to Succeed 1
Dr. Charles Browne, Meiji Gakuin University

2 Taking an English Language Curriculum Online 19
Steven Asquith, Phoebe Lyon, and Kathryn Jurns, Kanda University of International Studies

3 The Current State of MALL in Japanese Universities 30
Edo Forsythe, Hirosaki Gakuin University

4 Enhancing Teacher Education Through a Lesson Study Application Built with WordPress 47
James M. Hall, Kodai Tamura, and Shuhei Kumagai, Iwate University

5 Exploring EFL Student Use of Digital Backchannels During Collaborative Learning Activities 64
Peter Ilic, University of Aizu

6 Comparing Writing vs. Smartphone Tapping Speed 78
Bradford J. Lee, Fukui University of Technology

7 Law Undergraduates' Understanding and Appropriation of Arguments in Online Essay Writing Tutorials 91
Suman Luhach, Bennett University, India

8 A Comparison of the Affective Affordances of a Static and Interactive VR System on Learner FLA and Motivation 108
Koichi Shibata and James York, Tokyo Denki University

9 The Effect of Modality on Oral Task Performance in Voice, Video, and VR-based Environments 128
Hayato Tokutake, James York, and Hiroshi Nakayama, Tokyo Denki University

https://doi.org/10.37546/JALTSIG.CALL2020.1

1

The NGSL Project: Building Wordlists and Resources to help EFL Learners (and Teachers) to Succeed

Dr. Charles Browne, Meiji Gakuin University

Abstract

During my JALTCALL 2020 Plenary Address, I explained about the importance of high frequency and special purpose (SP) vocabulary for second language learners of English, and then went on to introduce our New General Service List Project, a collection of seven open-source, corpus-based word lists offering the highest coverage in each of their specific genres, as well as the large and growing number of free apps and online tools we have either developed or utilized to help learners, teachers, researchers and materials developers to better be able to utilize our lists. This chapter is a very brief summary of this project.

JATCALL2020の基調講演で、私は英語を第二言語として学ぶ学習者にとっての高頻出語彙と特殊目的の語彙の重要性を説明し、New General Service Listプロジェクトを紹介した。本プロジェクトは、ジャンルごとに広範な語彙を提供する７つのオープンソースを取りまとめているコーパスベースの単語リストや、そのリストを学習者、教員、研究者、教材開発者がよりよく利用できるようにするために開発・利用してきた無料アプリやオンラインツールを含んでいる。この章ではこのプロジェクトについて簡単にまとめる。

Keywords: **NGSL, vocabulary learning, corpus linguistics, NAWL**

A New General Service List

Back in 1953 Michael West published an incredibly useful list of high frequency words known as the General Service List or GSL (West, 1953). Based on more than two decades of pre-computer corpus research, input from other famous early 20th century researchers such as Harold Palmer and partially funded by the Carnegie Foundation and Rockefeller grants, the GSL was designed to be more than simply a list of high frequency words. The GSL went beyond that, combining both objective and subjective criteria to create a list of words that would be of general service to learners of English as a foreign language.

Although the list has helped teachers, students, and researchers for decades, it suffered from several problems including being based on a relatively small (2.5 million words) and dated corpus. Because language usage changes over time, it was clear that the GSL needed to be updated and in early 2013, the New General Service List or NGSL (Browne, 2013)

was published. The starting point for the NGSL was a carefully selected 273-million-word subsection of the 2-billion-word Cambridge English Corpus (CEC).

Table 1

New General Service List Corpus Overview

CEC Sub-Corpus	Tokens
Learner	38,219,480
Fiction	37,792,168
Journals	37,478,577
Magazines	37,329,846
Non-Fiction	35,443,408
Radio	28,882,717
Spoken	27,934,806
Documents	19,017,236
TV	11,515,296
Total	273,613,534

Like the GSL before it, the NGSL combines both quantitative criteria (frequency analysis of the corpus) with qualitative (the input and advice from vocabulary specialists and expert teachers). With over 600,000 words in the English language and most adult native speakers of English knowing at least 30,000 words, the 2800 words of the NGSL offers a surprisingly high 92% coverage for most general English books, newspapers and magazines, and even higher coverage for most TV shows and movies. Recent research (Iwamoto & Kramer, 2020) has shown that in addition to providing good coverage for general English in daily life, the NGSL is also very useful for test preparation in Japan, with the full 2800 word list providing 95.2% coverage of the National Center Test, and study of just the first 1000 most frequent NGSL words providing 98.1% coverage of the High School Entrance Exam. The NGSL and all of our word lists are available for free download and use under one of the least restrictive Creative Commons Licenses from the website **http://ww.newgeneralservicelist.org**.

Why Zipf's Law is So Important for Language Learners

There is a mathematical principle known as Zipf's law, which shows us that while on the one hand, the highest frequency words are very few in number yet occur with unbelievably frequency in our daily lives, the inverse is true for low frequency words, which are huge in number and are almost never met. Below is a summary of some of the frequency research. In it you can see that the first 8000 most frequent words of English represent 98% of the words most learners would ever meet, whereas the remaining 592,000 words of the English language would only give an additional 2% coverage.

Table 2

Zipf's Law in Frequency Data

Words	% known	# known	Researcher
1	7%	97/100	West (1953), Nation (1990)
10	25%	3/4	West (1953), Nation (1990)
100	50%	1/2	West (1953), Nation (1990)
1000	75%	1/4	West (1953), Engles (1968)
2200	80%	1/7	West (1953), Nation (1990)
2801	**92%**	**1/13**	**Browne et al (2013)**
5000	95%	1/20	Hirsch & Nation (1992)
8000	98%	1/50	Laufer (1992), Coady (1993)
600,000	100%	100/100	Oxford English Dictionary

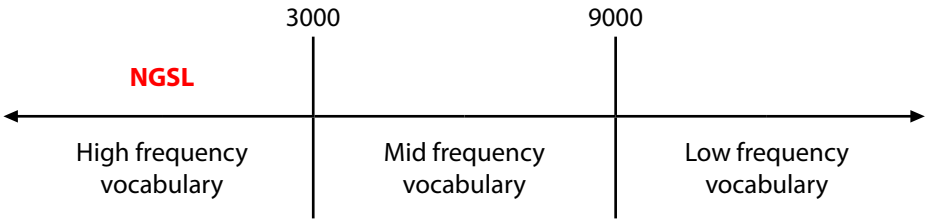
We know from the research on vocabulary thresholds that a minimum of 90% but preferably 95–98% of the words in a reading or listening text must be known in order to be able to understand it (Laufer, 1989, 1992), so what Zipf's law means for second language learners is that in order to reach these important coverage targets, it is absolutely essential for them to know all of the words in the NGSL since any significant gap in knowledge of these ultra-high frequency words would make it virtually mathematically impossible to achieve 90% or higher coverage.

Next Steps: Mid-Frequency and Special Purpose Vocabulary

So what is the next step after the NGSL? Perhaps the next most logical possibility is a focus on mid-frequency vocabulary. Though they occur with less frequency than words in the NGSL, they are much more likely to be encountered than low frequency words, and when studied systematically, can increase percent coverage of texts from the low 90s (offered by the NGSL) to the high 90s.

Table 3

Mid-Frequency Vocabulary



The problem is that it takes many years to learn the first 3000 most frequent words of English, more than 12 years of public school study in the case of Japanese EFL learners This

means that learning another 6000 mid-frequency words is not a practical target for most learners as it would take more than another 2 decades of study at the same rate of learning!

A potential shortcut for certain language learners is the study of English for Specific Purposes (ESP) vocabulary. Mastering the highest frequency ESP words within a specific genre can help learners to reach higher coverage figures much sooner than if they continue to study general English vocabulary at the mid-frequency level.

As the NGSL Project is very much about efficiency and creating shortcuts for learners, we thought it would make good pedagogic sense to create solid, corpus-based ESP word-lists as a second step for students with that particular goal (after mastering the core, foundational words of the NGSL). To date, we have created four additional ESP word lists, for the genres of academic English, TOEIC preparation, business English and fitness English.

A New Academic Word List

Many years after the original GSL was published, an Academic Word List (AWL) was created (Coxhead, 2000) to go along with the GSL as a 2nd step in vocabulary development for second language learners hoping to go on to do university level studies in English speaking countries. Although this special purpose (SP) vocabulary list was very well constructed and offered good coverage for academic texts, it was designed to work together with the original GSL rather than the NGSL, so the next step was to develop our own academic word list that would align better with the NGSL (i.e., each list represents the highest frequency words in that genre but with no overlapping words).

First published in May of 2013 (Browne et al., 2013) and available for free download at **http://www.newgeneralservicelist.org/nawl-new-academic-word-list**, the New Academic Word List (NAWL) is based on a 288 million word academic corpus (see the summary below in Table 4) and offers an average of 92% coverage of academic texts and lectures, when learned in conjunction with the NGSL.

Table 4

New Academic Word List Corpus Overview

Corpora	Size	Percent
CEC Academic	248 million	86.30%
Oral	3 million	1.10%
Textbooks	36 million	12.60%
Total	288 million	100%

CEC: Academic journals, non-fiction, student essays, academic discourse
Oral: MICASE (Michigan Corpus of Academic Spoken English), and BASE (British Academic Spoken English)
Textbooks: Corpus of 100s of top-selling academic textbooks

English for TOEIC

TOEIC is a high stakes English proficiency exam produced by English Testing Service (ETS) and is said to be a measure of business English ability. The 1200 word TOEIC Service List, also known as the TSL (Browne & Culligan, 2015a) is available for free download at **http://www.newgeneralservicelist.org/toeic-list**, and is based on a carefully constructed 1.5 million word corpus of various TOEIC preparation materials. At the time of release in 2015, we found that when learned in conjunction with the NGSL, the TSL offers 99% coverage of the TOEIC exam.

Below is a list of some of the TOEIC textbooks, practice tests, and TOEIC preparation materials included in the corpus:

600 Essential Words for the TOEIC – 3rd Edition
Achieve TOEIC Bridge
Barron's TOEIC Practice Exams
Collins' Practice Test for the TOEIC Test
Oxford practice tests for the TOEIC test
Oxford Preparation Course for the TOEIC test
Pass the TOEIC Test Advanced
Pass the TOEIC Test Intermediate
Practice Examinations for the TOEIC Test
Tactics for TOEIC Listening and Reading Tests

In 2018, two years after the publication of our TSL, the format of the TOEIC test underwent a few changes in format and question type, so in order to see if the TSL would still offer good coverage of the new format, in 2019, we created a small corpus from the 8 official TOEIC practice tests and found that the combination of learning the NGSL and TSL provides 98.5% coverage of the latest version of the TOEIC exam.

Business English

Although TOEIC is said to be a test of business English, anecdotal evidence such as the failure of the TOEIC initiative at Rakuten, as well as independent research that has been critical and unsupportive of ETS's claims that it is good measure of business English ability (Childs, 1995; Hirai, 2002) led us to feel that separate from our TSL word list, which was designed help students improve their scores on a high stakes exam, we should also release a business English vocabulary list based on a corpus that is more representative of the business English that EFL and ESL learners were likely to encounter in the real world.

The 1700 word Business Service List, also known as the BSL (Browne & Culligan, 2015b) is available for free download at: **http://www.newgeneralservicelist.org/bsl-business-service-list** and is based on a carefully constructed 64 million word corpus of business English texts as follows:

Table 5

Business English List Corpus Overview

Business Corpus	Size
BNC	8,386,492
Internet	9,199,615
Magazines	25,085,121
Textbooks	21,980,494
Total	64,651,722

Fitness English

Very recently (Browne & Culligan, 2020a), we published a list of high frequency fitness English words known as the Fitness English List (FEL). Originally developed for a chain of English-medium gyms in Japan, we have been given permission to share our research and NGSL-compatible word list with the general public (the chain of gyms is using a different version of the FEL that does not include the NGSL). Based on a 9.7-million-word corpus of various fitness-related texts, websites and YouTube channels, the 600 words of the FEL gives up to 98% coverage of common fitness English when combined with the 2800 words of the NGSL. It is available for free download and use from **http://www.newgeneralservicelist.org/fitness-english-list**. A summary of the corpus appears below:

Table 6

Fitness English Corpus Overview

Sub-Corpus	Size (number of words)	Percent of Total Corpus
Fitness Certifications	176,037	1.8%
Flexibility Training	5,273,464	54.0%
Functional Training	1,470,372	15.1%
Nutrition-related	528,381	5.4%
Recovery-related	330,827	3.4%
Strength Training	1,979,129	20.3%
Total	9,758,210	100%

New Dolch List: Paving the Way for Young Learners of English

Though the NGSL is an extremely useful first step for adult learners of English, the type of English that young learners of English are exposed to is significantly different, than that of adults, so in the summer of 2020 the New Dolch List, or NDL (Browne & Culligan, 2020b) was published and is available for free download at

http://www.newgeneralservicelist.org/new-dolch-list. The NDL was developed with the goal of creating a reliable and valid corpus-based list of high frequency English words important for young second language learners to be able to successfully interact with EFL learning materials, other EFL learners, as well as popular children's TV shows and children's picture books.

The NDL is a significant update of the original Dolch list, which was published back in 1936 by Edward William Dolch (Dolch, 1936). That list contained 220 *sight words,* words which need to be quickly and easily recognized to achieve reading fluency in English as well as an additional list of 95 important nouns. It has been argued that up to 70% of all words used in schoolbooks, library books, newspapers, and magazines are a part of the Dolch basic sight word vocabulary. Though quite dated, the Dolch lists are still widely assigned for memorization in American elementary schools and used in ESL and EFL settings and materials around the world.

That said, a wordlist is only as good as the corpus it is based on. Like West's 1953 General Service List (GSL), which was replaced by the New General Service List (Browne, 2013), the Dolch 1936 list has often been criticized for being based on outdated resources. The English language changes and evolves over time and an update to the 90-year-old Dolch list was long overdue.

The original Dolch list was based on a corpus and word lists designed only for native speakers of English attending primary school in English speaking countries. English is now taught and studied as a second language in countries around the globe and the original Dolch list was not designed with these needs in mind. EFL learners do not get as much input in English as those living in English speaking countries and the sources of input are more limited. EFL textbooks, children's songs, children's textbooks and children's TV shows are usually the primary sources of input for most EFL learners and we have created a corpus of such materials to generate a list of the most important words for EFL learners in the hopes that this would be a valuable asset to EFL teachers, students, textbook authors and educational software developers around the world. The NDL is based on a carefully selected, 2.5-million-word corpus of children's reading and listening materials as follows:

Table 7

New Dolch Children's Corpus Overview

Category	Total
Children's Graded Readers	991,104
Elementary EFL Textbooks (MEXT)	4,623
Middle School Textbooks	322,169
Picture Books (for L1)	13,347
EFL Textbooks	575,976
Top Kids YouTube Videos	548,172
New Dolch Corpus Total	2,455,391

The 315 words of the original Dolch word list (220 sight words plus 95 nouns) were

said to offer up to 70% coverage, though there do not seem to be any empirical papers to verify this claim. In looking at the NDL, it can be seen that at 315 words, coverage has already surpassed this, giving 78% coverage:

Table 8

Preliminary Coverage Data for the NDL

305	877	under	1593	Functors	−5.341	0.0004	0.77	NGSL0193	1–NGSL
306	876	end	4961	Lexeme	−3.725.	0.0004	0.77	NGSL0152	1–NGSL
307	876	repeat	62177	Lexeme	−3.055	0.0004	0.77	NGSL1251	1–NGSL
308	875	clothes	6926	Lexeme	−2.160	0.0004	0.77	NGSL0955	1–NGSL
309	874	often	19308	Lexeme	−4.806	0.0004	0.77	NGSL0288	1–NGSL
310	871	board	31366	Lexeme	−0.164	0.0004	0.77	NGSL0741	1–NGSL
311	870	ball	405	Lexeme	−2.841	0.0004	0.77	NGSL1199	1–NGSL
312	863	visit	5184	Lexeme	−3.934	0.0004	0.77	NGSL0333	1–NGSL
313	862	along	6501	Functors	−3.589	0.0004	0.77	NGSL0388	1–NGSL
314	854	everyone	7367	Lexeme	−4.087	0.0004	0.77	NGSL0650	1–NGSL
315	849	light	2693	Lexeme	−5.697	0.0004	0.78	NGSL0349	1–NGSL
316	846	cold	1909	Lexeme	−3.593	0.0004	0.78	NGSL0866	1–NGSL
317	844	die	18972	Lexeme	−4.370	0.0004	0.78	NGSL0454	1–NGSL
318	841	own	19090	Lexeme	−3.671	0.0004	0.78	NGSL0129	1–NGSL
319	841	train	2814	Lexeme	−4.359	0.0004	0.78	NGSL0289	1–NGSL
320	840	side	4814	Lexeme	−3.036	0.0004	0.78	NGSL0276	1–NGSL
321	840	sleep	2708	Lexeme	−5.491	0.0004	0.78	NGSL0826	1–NGSL

Unfortunately, SLA research shows that 78% is not nearly enough. Teachers often tell their students that when they come across an unknown word in a text that they should not stop and look it up in a dictionary but rather to guess the meaning from context. Research on vocabulary thresholds show that for this to be achieved, learners need to know a minimum of 90% of the words on the page, but preferably 95–98% (Laufer, 1989; 1992).

Therefore, in order to help students, schools and teachers get closer to the coverage levels needed to read and guess unknown words from context more easily, the NDL list was extended to 875 words, which offers 90% coverage for most EFL materials for young learners.

The NDL wordlist can be downloaded from the above link in a variety of formats including with just headwords, lemmatized for teaching purposes, and lemmatized for research purposes. We also have provided definitions in simple English for all words. In the near future, the NDL wordlist will also be added to free flashcard sites like Quizlet (**http://quizlet.com**), Memrise (**http://memrise.com/**), Word-Learner (our own flashcard app), and the free Online Graded Text Editing Tool (OGTE) which can be found at **https://www.er-central.com/ogte/**.

Spoken English

In response to several requests for a list of high frequency spoken English words, the three spoken subsections of the NGSL corpus were analyzed and in late 2013, Browne and Culligan published the NGSL-S (Browne & Culligan, 2013). The 1.0 version contained 822 words which provided 89% coverage for spoken English. In 2016, the 1.1 version of the list was published, with slightly better coverage (718 words to reach 90%). In October 2017 the 1.2 version of the NGSL was released. This list is 721 words and provides up to 90% coverage for unscripted spoken English (explained in Table 9).

Table 9

New General Service List-Spoken Overview

CEC Sub-Corpus	Tokens
Learner	38,219,480
Fiction	37,792,168
Journals	37,478,577
Magazines	37,329,846
Non-Fiction	35,443,408
Radio	28,882,717
Spoken	27,934,806
Documents	19,017,236
TV	11,515,296
Total	273,613,534

The Tools

We have created and made use of a large and growing number of free online tools to help teachers, students, researchers, and content creators to be able to utilize our word lists.

Wordlists & Definitions

The first set of tools that should be mentioned is the wordlists themselves. For each wordlist we have tried to make several downloadable versions available, including in alphabetical order, frequency order, lemmatized for teaching purposes, lemmatized for research purposes and also with basic statistical data from our analysis.

In order to make use of these wordlists in the classrooms as well as via many online learning platforms, we realized that teachers and learners would need to have definitions for each word and we did our best to write and provide definitions in simple English for all wordlists. Please note that these definitions were written very quickly, and we sometimes needed to pay people to help us write them. This means that although the downloadable definitions are original and free for use as people wish, but they have not been carefully edited. Please feel free to update or change the definitions as necessary. If there are any glaring errors,

please let us know and we will update the definition files available for download. Please also note that if you want very carefully edited definitions or sample sentences for all of our wordlists, these are available as part of our Word-Learner app (which contains flashcards, games, a 90k learner dictionary in 12 languages and an LMS).

One of the newest and most interesting resources that has been created is the interactive glossaries which include direct links to an in-context video concordance to help learners hear pronunciation of the word in various authentic contexts, something that is also great for helping to develop collocational knowledge. The NGSL Glossary can be accessed from **http://www.newgeneralservicelist.org/ngsl-glossary**. After you see one clip for the word you are interested in, click the forward button – most words have thousands of examples lined up to see. Below is a screenshot of a small section of the NGSL glossary. If you click on the HTML pronunciation link for the word *absolutely* you are brought to a cliplist player that will show more than 85,000 authentic video clips that use the word absolutely in various contexts. The first clip is shown in Figure 1 below:

Figure 1

Screen Shots 1 and 2: Interactive NGSL Glossary and Video Concordancer

NGSL Glossary

Word	PoS	FpM	Pronunciation	Definition
a	art	27872	/ɛɪ/	one (of something)
abandon	verb	37	/ə•bæn•dən/	to leave someone you should stay to take care of
ability	n	123	/ə•bɪ•lɪ•tɪ/	the power, money, or skill to do something
able	adj	361	/ɛɪ•bl/	having the ability to do something
abortion	n	24	/ə•boə•ʃən/	a medical operation to end a pregnancy
about	prep	2967	/ə•baʊt/	of (a subject); concerning
above	adv	228	/ə•bʌv/	in a higher place than something else
abroad	adv	39	/ə•broə•d/	in or to a foreign country
absence	n	43	/æb•səns/	fact of not being present, or not existing
absolute	adj	27	/æb•sə•lut/	complete and total
absolutely	adv	84	/æb•sə•lut•lɪ/	completely; totally; very
abstract	adj	21	/æb•strækt/	concerning something that doesn't exist physically
abuse	verb	52	/ə•bjuz/	to use in an inappropriate or excessive way
academic	n	43	/æ•kə•dɛ•mɪk/	educator who works at a college or university
accept	verb	172	/ək•sɛpt/	to agree to receive or take something offered
acceptable	adj	26	/ək•sɛp•tə•bl/	satisfactory; good enough
access	n	119	/æk•sɛs/	way to enter a place, e.g. A station or stadium

Tests

The New General Service List Test (NGSLT) is a diagnostic test of written receptive knowledge of the New General Service List (NGSL) that was developed by Phil Bennett and Tim Stoekel (Bennet & Stoekel, 2013; Stoekel & Bennet, 2015). The NGSLT is comprised of 100 items, 20 for each of five approximately 560-word bands of the NGSL. Both the New Academic Word List Test (NAWLT) and the NGSLT have been designed to the same specifications as the Vocabulary Size Test (VST) (Nation & Beglar, 2007). Each item provides the target word followed by a short sentence containing this word in a natural yet non-defining context. The test shows very high reliability (over .90), and was designed to help place students in approximately the right frequency band for study of NGSL words via the Quizlet word stacks or our free NGSL apps (NGSL Builder, described below). The test, answer sheet, and explanation can be downloaded directly from a dedicated webpage on our NGSL website (**http://www.newgeneralservicelist.org/ngsl-levels-test**).

The NAWLT (Bennet & Stoekel, 2013; Stoekel & Bennet, 2015) is a diagnostic test of written receptive knowledge of the New Academic Word List (NAWL). The NAWL is a list of 963 words that appear frequently in academic discourse and that are not part of the New General Service List (Browne et al., 2013). The NAWLT is comprised of 40 items, 20 from

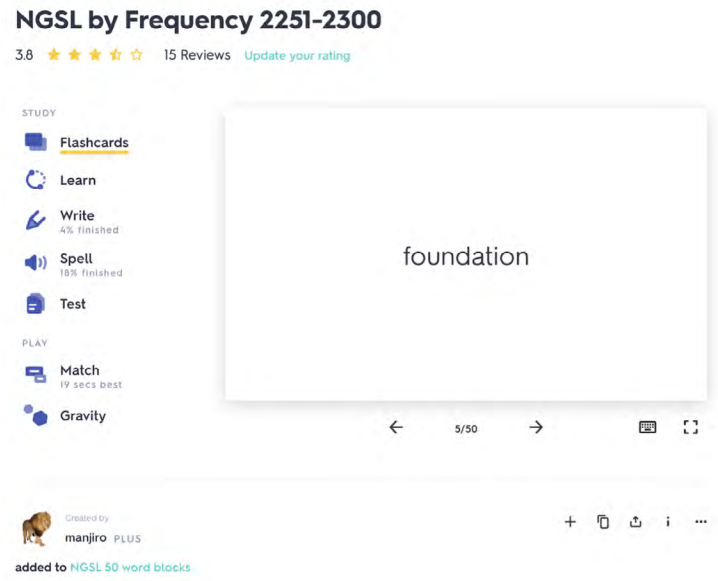
each of two frequency-determined bands of the NAWL. This sampling rate is approximately equal to that of the NGSLT by the same authors. Sampling from bands of around 500 words is useful for two reasons. First, it allows for more precise feedback than sampling from 1000-word bands, which has been the norm in recent years. Second, it was felt that approximately 500 words is a reasonable learning goal for a semester of study, especially considering that some of those words will already be known to learners. The test was designed to help place students in approximately the right frequency band for study of NAWL words via the Quizlet word stacks or the free NAWL apps (NAWL Builder, described below). The test, answer sheet, and explanation can be downloaded directly from this page via the same link as was given above for the NGSLT.

Flashcards

One of the first free flashcard websites you can use to study our wordlists is Quizlet. More than just flashcards, Quizlet is a site that allows learners to study wordstacks via 5 study and 2 game-like activities.

Figure 2

Screenshot of Quizlet

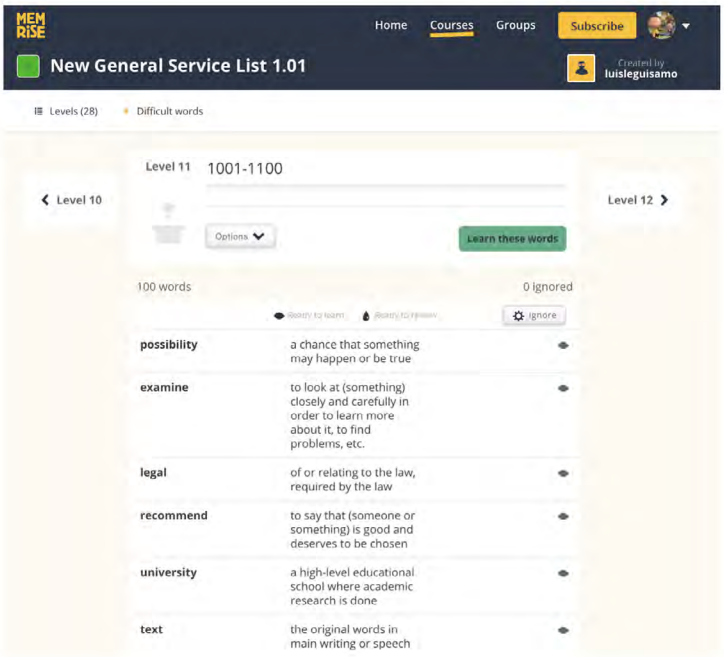


We have created stacks for the wordlists in various sizes to accommodate different purposes. For example, the NGSL can be studied in 560 word blocks which matches the categories and results from the NGSLT, and is said to be a good number of words to set as a target for a full semester or year of study, or it can be studied in more granular 100 word blocks or 50 word blocks. The easiest way to access these stacks is from the direct links provided on the NGSLT website (**http://www.newgeneralservicelist.org/ngsl-for-quizlet-soon**).

Another popular free flashcard website that we have made use of is Memrise. Like Quizlet, Memrise allows you to upload any list of words and definitions for study.

Figure 3

Screenshot of Memrise

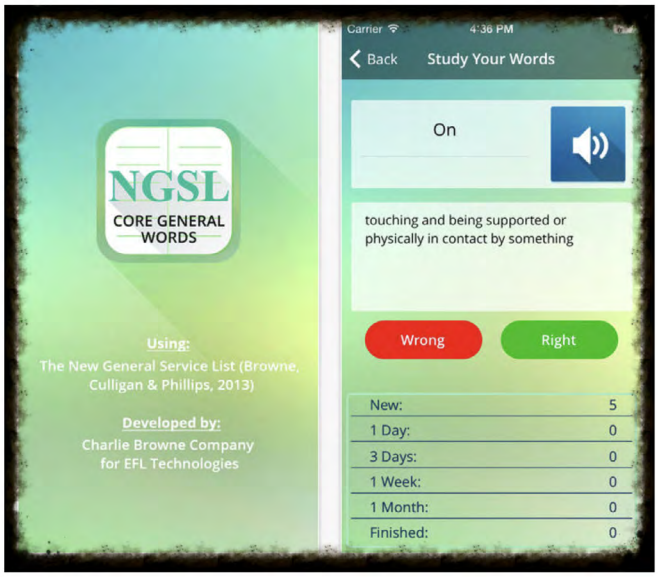


Although it does not offer the many different learning modalities of Quizlet, Memrise does offer something that Quizlet does not, and that is the ability for students to modify and personalize their flashcards, adding pictures and text to help them better remember their new words. You can access the NGSL wordstacks we have created via the direct links on the NGSL website at (**http://www.newgeneralservicelist.org/ngsl-for-quizlet-soon**).

I have also created a number of free smartphone applications to help with the learning of our wordlists, all of which make use of spaced-repetition algorithms to enhance the speed and depth of learning. The earliest flashcard apps I made are called NGSL Builder and NAWL Builder. You can find English and Japanese versions of the app in both the iOS and Android app stores. Figure 4 below includes some screen shots from NGSL Builder:

Figure 4

Screenshot of the NGSL Builder App

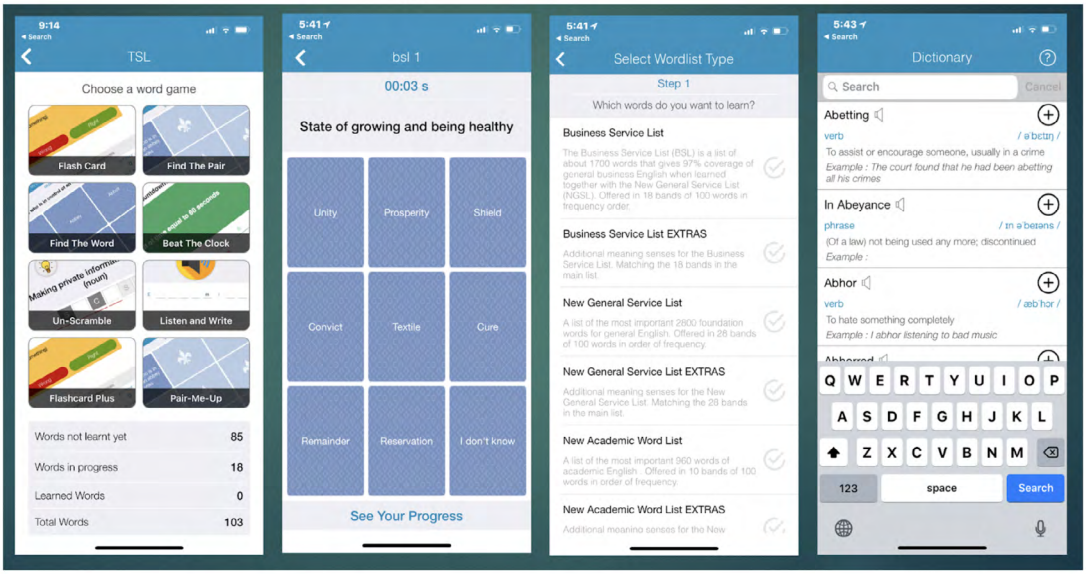


As a side note, I have also created free apps for the original GLS and AWL word lists as well. All of the above apps can be downloaded from the iOS and Android app stores or also accessed from our dedicated web page here (**http://www.newgeneralservicelist.org/ngslnawl-iphone-apps**)

More recently, Rob Waring and I have released a more advanced learning app called Word-Learner. Like NGSL Builder, Word-Learner also uses spaced-repetition algorithms to help improve learning, but also uses principles of gamification to make the learning more fun. In addition to flashcards, the app allows learners to work with their word stacks using flashcards as well as seven other game-like activities, adding a point system to track their progress. Word-Learner also provides free access to our 90,000-word proprietary learner dictionary which gives carefully edited definitions and sample sentences in easy English as well as eight other languages including Japanese. The app contains all NGSL word lists as well as several Cambridge wordlists, the Oxford 3000, SAt and CEFR-J (see Figure 5). It also provides teachers with a learner-management system (LMS) and has the ability to give students placement and progress tests. As with NGSL Builder, the app can be found on both the iOS and Android app stores, or you can get more information from the dedicated app web page here (**http://www.newgeneralservicelist.org/ngslnawl-iphone-apps**).

Figure 5

Screenshot of the Word-Learner app



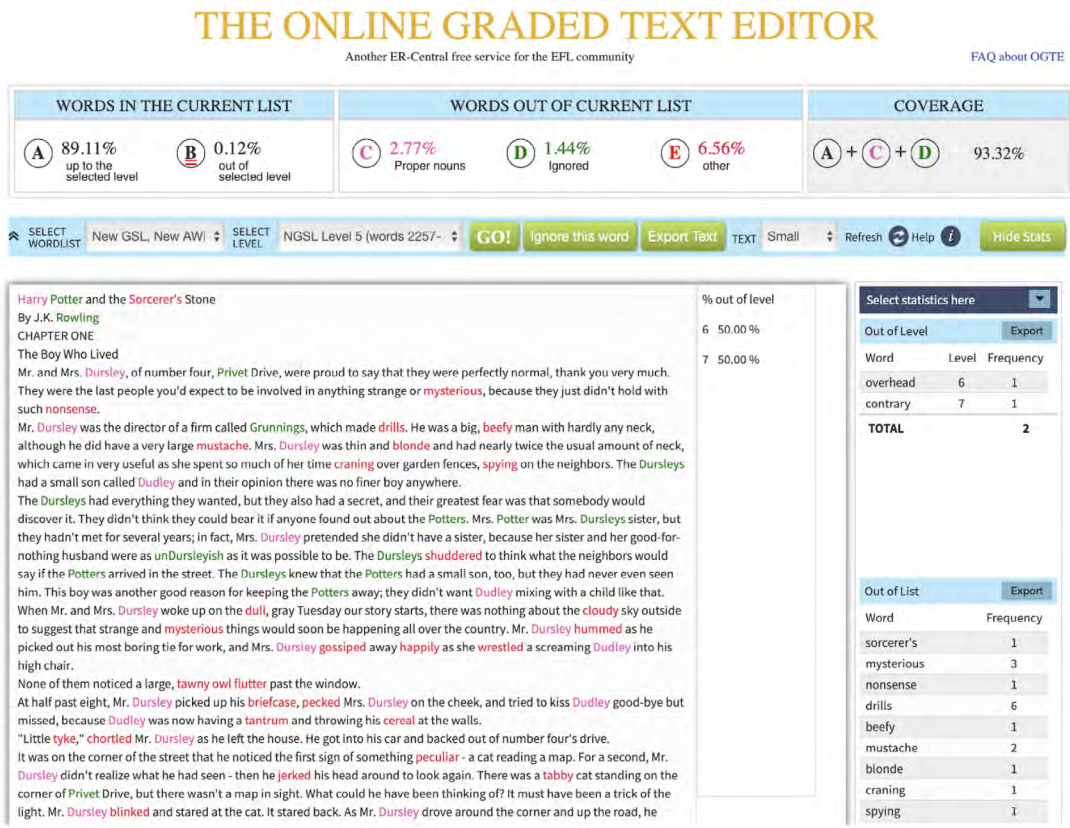
Text Analysis Tools

One of the important principles of doing graded, extensive reading or listening with your students is to make sure they are getting *Comprehensible Input* at their *i*+1 level (Krashen & Terrell, 1983; Krashen, 1985). Many researchers have quantified this *i*+1 level to be the point where students know at least 95-98% of the words on the page (Laufer, 1992; Nation, 2000). One of the main purposes of the NGSL Project is to get students to this level as quickly as possible. Our word lists and learning tools are there to help students to master the highest frequency words systematically and efficiently. But what about the materials or content side of the equation? How does a teacher decide if authentic reading or listening materials are at an appropriate level for their students? How do they go about modifying and simplifying authentic materials to the level of their students? How do they create or write original materials at the students' level? This is the main purpose of the free Online Graded Text Editor Tool (OGTE) (Browne & Waring, 2015), which can be accessed from **http://www.er-central.com/ogte/**.

With OGTE, users can choose one of our word lists to see how much coverage it provides of the text you want to use with your students. Color coding helps to make the output easier to understand. For example, in Figure 6 below, it is clear that knowing the NGSL would provide 93.32% coverage for the first chapter of *Harry Potter*. Words in black are the NGSL words with pink words being proper nouns – that can be ignored since students usually do not have difficulty with place and people names – and words in red being the off-list words that are potentially difficult for the students.

Figure 6

Screenshot of the NGSL Coverage for *Harry Potter* Using OGTE Tool



The tool allows teachers to make judgments about the suitability of a text for their students (if they know the NGSL, *Harry Potter* is fairly close to being at their *i*+1 level), and also can help teachers to be able to quickly simplify texts that are too difficult for their students. In the above text, if a user wanted to get it closer to the 95–98% threshold you could try removing a few difficult words, for example *mysterious* or *beefy*, or replacing difficult words for simpler ones: for example, replacing *craning* with *looking*.

In addition to our own OGTE tool, we have made our NGSL wordlists available for other similar tools such as the VocabProfile tool by Tom Cobb (**https://www.lextutor.ca/vp/comp/**), or AntWordProfiler by Laurence Anthony (**www.laurenceanthony.net/software/antwordprofiler/**).

Conclusion

Michael West began work on the original GSL in 1936 and did not publish his final version of the list for 17 years until 1953. The first version of the NGSL was published in 2013. In the ensuing 7 years the NGSL list has been updated and improved several times,

published six additional lists, and a growing multitude of online tools to better utilize these lists in and out of the classroom, in research and in content creation. We figure that gives us at least another 10 years to work out the kinks in the various parts of the NGSL Project. Our lists and tools are imperfect but we have made them all open-source and free, and it is our strong intention to continue to publish additional lists and tools while doing our best to improve the existing ones, all with the overriding goal of making the lives of EFL students, teachers, researchers and content creators a little better.

References

Bennett, P., & Stoeckel, T. (2013). Developing equivalent forms of a test of general and academic vocabulary. In N. Sonda & A. Krause (Eds.), *JALT2012 conference proceedings*. JALT.

Browne, C. (2013). The New General Service List: Celebrating 60 years of vocabulary learning. *The Language Teacher, 37*(4), 13–15.

Browne, C., & Culligan, B. (2013). The New General Service List – Spoken. **http://www.newgeneralservicelist.org/ngsls**.

Browne, C., & Culligan, B. (2015a). TOEIC Service List. **http://www.newgeneralservicelist.org/toeic-list**.

Browne, C., & Culligan, B. (2015b). Business Service List. **http://www.newgeneralservicelist.org/bsl-business-service-list**.

Browne, C., & Culligan, B. (2020a). The Fitness English List. **http://www.newgeneralservicelist.org/fitness-english-list**.

Browne, C., & Culligan, B. (2020b). The New Dolch List. **http://www.newgeneralservicelist.org/new-dolch-list**.

Browne, C., & Waring, R. (2015). The Online Graded Text Editor. **http://www.er-central.com/ogte/**.

Browne, C., Culligan, B., & Phillips, J. (2013). New General Service List. **http://www.newgeneralservicelist.org/nawl-new-academic-word-list**.

Childs, M. (1995). Good and bad uses of TOEIC by Japanese companies. In Brown & Yamashita (Eds.), *Language testing in Japan* (pp. 66–75). JALT.

Coady, J. (1993). Research on ESL/EFL vocabulary acquisition: Putting it in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning,* (pp. 3–23), Ablex.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*(2), 213–238.

Dolch, E.W. (1936). A basic sight vocabulary. *The Elementary School Journal, 36*(6), 456–460.

Engels, L.K. (1968). The fallacy of word-counts. *IRAL, 6*(2), 213–231.

Hirai, M. (2002). Correlations between active skill and passive skill test scores. *Shiken: JALT Testing & Evaluation SIG Newsletter, 6*(3), 2–8.

Hirsh, D., & Nation, I.S.P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language, 8*(2), 689–696.

Iwamoto, M., & Kramer, B. (2020). NGSL Coverage of the SHS Entrance Exam and the National Center Test. *The Language Teacher, 44*(5), 3–7.

Krashen, S.D. (1985). *The input hypothesis: Issues and implications.* Longman.
Krashen, S.D., & Terrell, T.D. (1983). *The natural approach: Language acquisition in the classroom.* Prentice Hall Europe.
Laufer, B. (1989). What percentage of text lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316–323). Multilingual Matters.
Laufer, B. (1992). How much lexis is necessary for reading comprehension? In H. Béjoint & P. Arnaud (Eds.), *Vocabulary and applied linguistics* (126–132). Macmillan.
Nation, P. (2000). *Learning vocabulary in another language.* Cambridge University Press.
Nation, I.S.P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher, 31*(7), 9–13.
Stoeckel T., & Bennett, P. (2015). A test of the New General Service List. *Vocabulary Learning and Instruction, 4*(1), 18.
West, M. (1953). *A General Service List of English Words.* Longman, Green, and Co.

Author's Bio

Dr. Charles Browne is a Professor of TESOL and Applied Linguistics and head of the EFL Teacher Training Program at Meiji Gakuin University in Tokyo Japan. Specialist in corpus linguistics, vocabulary and reading, especially as they apply to online learning environments. Creator of many free, open-source vocabulary lists, learning apps, text analysis tools, extensive reading websites and more.

https://doi.org/10.37546/JALTSIG.CALL2020.2

2

Taking an English Language Curriculum Online

Steven Asquith, Phoebe Lyon, and Kathryn Jurns, Kanda University of International Studies

Abstract

Whilst online courses have become ever more prevalent in the educational field in recent decades, their efficacy is still debated, especially with respect to language communication classes, which traditionally entail human interactions. Although there has been previous, limited research conducted into online learning options at the same tertiary institution (Mynard & Murphy, 2012), online courses were still far from ubiquitous in the Japanese landscape of tertiary language education curricula at the time of this study. However, adoption of online courses would allow for the expansion of current programs, enabling non-traditional students access to equal education opportunities as well as offering institutions a practical alternative to having to cancel lessons owing to unforeseen circumstances. This study investigates whether online lessons of the core curriculum classes offered in a language-focused Japanese university context, and which were based upon the regular in-class course content, offer comparable value to the regular classroom-based lessons in terms of students' perceptions of the delivery, structure, and usefulness of the lessons. By providing practical descriptions of how the traditional classroom lessons were adapted and delivered online, as well as quantitative feedback comparing the students' perceptions of the online versus classroom lessons, the paper will highlight the challenges of creating online course content, considering both instructors' and students' viewpoints.

昨今の教育現場において、オンライン授業は広く取り入れられるようになったが、外国語コミュニケーション科目に代表される対話に重きを置く授業では、オンライン化によって具体的にどのような効果が得られるのかまだ論議の余地がある。過去に当学でオンライン授業に関する小規模な研究が行われたが (Mynard & Murphy, 2012) 講義のオンライン化はその当時、日本の大学の語学教育カリキュラム上で全く普及していなかった。しかしながらオンライン授業を採用することにより、社会人を対象に教育の門戸が開かれ、やむを得ない事情により休校・休講になった授業の振替処置がとりやすくなるなど、現存の語学プログラムの更なる発展を期待できる。この研究では、日本の一語学大学において正規の対面授業と同一の全学共通教育カリキュラムを用いたオンライン授業が、対面授業と比較した場合、学生が講義の進行、構成、有用性において同等の価値を得られたと感じるか調査する。従来の対面授業がオンライン化に至るまでの実用例をまじえ、学生を対象にした対面・オンライン授業への意識調査と共に紹介しながら、本論文では、教員・学生双方の視点からオンライン授業構成の抱える今後の課題に焦点をあてる。

Keywords: **online lesson, in-person lesson, blended learning, student perceptions**

キーワード: オンライン授業、対面授業、ブレンド型学習、学生の意識

As online learning opportunities have become more common in recent decades with the advancement of technology, and even more so in 2020 with the recent campus closures of educational institutions due to COVID-19, the fact remains that there is still considerable debate about whether online offerings are able to produce the same desired educational outcomes as in-person lessons. Whilst online learning environments have proven effective at accommodating students' different learning and social styles (Jeschofnig & Jeschofnig, 2011) as well as catering to non-traditional students, it is difficult to overlook the obvious drawback of a lack of face-to-face interaction. This study was conducted at a language-focused university in Japan. Being that the basis of lessons is on communication, this paper will draw on quantitative data and qualitative testimony to investigate whether the lessons offered online were able to offer comparable value based upon student and teachers' perceptions to the in-person, classroom-based lessons traditionally offered. Value is defined both quantitatively through the evaluative items' ratings and qualitatively through the perceptions of students and teachers.

Literature Review

This study is an analysis of students' perceptions of the value of online lessons in comparison to in-person lessons with respect to the language skills used, the amount of interaction occurring, and the lessons comparative perceived value. At the time of the study, the university did not offer online courses or lessons for students. However, a previous study conducted by Mynard and Murphy (2012), whereby students completed 100% of their tasks online in an "experimental day", showed that whilst students had found the online activities both useful and convenient, they had expressed concerns regarding possible limitations in technology and how these might negatively impact their experience. Of note, the purpose of this study was not to replace the traditional classroom setting, but rather to investigate the possibility of providing a blended learning environment to help accommodate a temporary change of schedule in the first semester of the 2020 academic year, resulting in 13 weeks on campus instead of the usual 15. Although the semester would be shortened, the amount of material that needed to be covered would remain the same. Thus, a blended learning approach, which involves delivering instruction "through a combination of physical and virtual instruction" and which aims to combine the best features of both (Garrison & Vaughan, 2011), was one option proposed by the university. If implemented, the online lessons were to be interspersed amongst the in-person lessons during the semester.

Fortunately, research indicates that blended learning courses improve learning outcomes for students (Garnham & Kaleta, 2002; Twigg, 2003) and have led to higher average scores (Kenney & Newcombe, 2011), and increased course completion rates in some cases (Garrison & Kanuka, 2004). An important consideration when adapting an in-person course to an online environment is ensuring that course goals and objectives are successfully met (Koszalka & Ganesan, 2004). As such, the online lessons that the teachers created were

designed with the key learning outcomes in mind, incorporating a variety of tasks that are typically found in in-person lessons across the curriculum, such as listening, reading, speaking, and writing skills which become progressively more challenging as students advance from first to second year. In addition to a focus on the four skills, of which some have more or less of an emphasis depending on the course, there are the following nine overarching outcomes: audiovisual analysis, awareness of self as learner, textual awareness and control, criticality and interpretation, textual fluency, interactive capacity, interpretation and expression of multimodal meaning, lexico-grammatical control, and intercultural capacity. These outcomes are founded in the multiliteracies approach implemented throughout the department and are described in detail in Johnson et al. (2016). In an effort to preserve a level of interactivity common in in-person lessons, teachers included collaborative discussion tasks within the online lesson materials. Whilst the nature of online/asynchronous discussions allows students to access the discussion at different times and therefore better enables them to control the pace of interaction (Huang, 2000), it is also promising that interactions in online contexts have been found to be highly valued in many studies (Bollinger, 2017; Li, 2015; de Freitas et al., 2015; Wells, 1999).

It is also important to consider that "technology quality, online tools and face-to-face support are predictors of learner satisfaction" (Kintu et al., 2017, p. 17). Due to the introduction of iPads into classes in 2014, which all students own and use in their lessons on a regular basis, a logical assumption would be that students would feel less apprehension towards technology in the classroom than they had in the Mynard and Murphy (2012) study. However, it should be noted that the students were not familiar with the Canvas learning management system (LMS) that was used for the purpose of the online lessons created during this study. Many students may have had the experience of using alternative LMSs such as Google Classroom or Schoology; however, this was only in the context of regular in-person lessons. For students participating in this study, this was the first time to take an online lesson at the institution and therefore they were not accustomed to having a lesson without a teacher present to offer support.

Methodology

The purpose of this study was to ascertain students' perceptions of the value of a one-off, online lesson, versus traditional, in-person lessons in the context of an international university. To do this, seven researchers created online lessons for the six core-courses of the English Language Institute (ELI) curriculum taught during students' first two years at the university. The courses are Freshman English and Foundational Literacies, both taught at the first-year level and which focus on listening and speaking, and reading and writing respectively. Media English (ME), English for International Communication 2 (EIC2), Academic Literacies: Reading (AR), and Academic Literacies: Writing (AW) are taught in the second year. ME and EIC2 focus on listening and speaking, while AR and AW focus on their namesakes.

Students who participated recorded their perceptions of online and in-person lessons in a survey as a basis for comparison of the two types of lessons. Both lesson types were conducted during their regularly scheduled lesson times. Each online lesson was given to two

classes for each of the six courses. In total, data from 124 students were analyzed. This figure was achieved by removing incomplete or duplicate responses and only selecting responses for which students completed both the in-person and online surveys. This resulted in a total of 124 pairs of responses (N = 124). The data was analyzed using a repeated measures analysis of variance (rANOVA) to measure if results were statistically significant. This measure was chosen because of the counterbalanced, within-subjects design of the study.

The online lessons were constructed using the LMS, Canvas, and were designed to reflect typical learning outcomes of a regular, 90-minute, in-person lesson. The lesson content creators included tasks incorporating listening, analysis, writing, and/or collaborative discussion practice. Additionally, in an attempt to better replicate a traditional lesson, many teachers included online individual response tasks and peer-to-peer discussions. It should be noted that the free version of Canvas was used, which is more limited in functionality than the paid version.

The administration of the online lessons was fairly straightforward. Prior to the lesson, students were oriented to Canvas. For some, this included an orientation lesson, but for others, there was a more basic walkthrough of the functions they would be using. On the experimental day, students were asked to come to their regularly scheduled lesson times. At the commencement of the lesson that was to be completed online, participants exited the classroom and went to a common area to complete the online lesson. They were told to communicate using online methods if they encountered any issues. Finally, students returned to the classroom 5–10 minutes before the end of the lesson to complete the survey for the study. While the study was meant to replicate a student's experience of doing an online lesson, the lessons were administered during regularly scheduled lesson times to assuage the workload required of the participants.

For the study, a quasi-experimental, within-subjects design with two conditions was used. The instrument, a survey, focused on the delivery, structure, and perceived value of the lesson as opposed to the content. Thus, the online lessons used similar processes to the traditional in-person lessons, just with different content. This involved each lesson creator adapting existing course materials to be delivered online. In order to improve reliability, a mixed-method research design was implemented to also gather feedback on the lessons through student comments. Participants were asked to complete the survey at two different points, once after an in-person lesson and once after an online lesson. To counterbalance the study, one class of the two participating classes from each course first completed the survey after the in-person lesson, before completing it again after the online lesson. The alternate class did the opposite, completing their first survey after the online lesson followed by the in-person lesson and survey. In both cases there was a rest period of one week between the two classes. The purpose of implementing this rest period was to give them time between each type of lesson; the hope being that they would rate the in-class lesson on its own merits and not rate it in comparison to the online lesson and vice versa.

The data collected were used to analyze students' perceptions of in-class versus online lessons in an attempt to understand the following three questions:

1. What language skills did students feel they were using in each lesson type (online and in-person)?

2. How much interaction occurred between classmates and with the classroom teacher in each lesson type?
3. How did students evaluate the level, usefulness, use of technology, ease to follow, and level of interest in each lesson type?

Results

Quantitative Analysis of Students' Perceptions of the Online Compared to In-person Lessons

Student perceptions of the online materials compared to traditional in-class lessons were recorded based upon three criteria: 1) Perceptions of the skills used, 2) Perceptions of the amount of interaction taking place, 3) Perceptions of level, usefulness, ease to follow, technology use, and interest. Participant responses (N = 124) were analyzed once incomplete, duplicate, and single responses were removed. A repeated measures analysis of variance (rANOVA) was used to analyze the data for statistical significance as this best suited the counterbalanced, within-subjects design of the study.

Skills and Interaction

Students' perceptions of the extent to which each skill was used during the lessons were calculated by asking them to input percentage values based on the question "During the lesson, how much, in terms of a percentage, did you use each of the following language skills?". Therefore, for instance, if each skill were being used equally all four mean values would be 25%.

Table 1

Descriptive Statistics of Participant Perceptions of Language Skills Use

Skill	Condition	*N*	Mean	*SD*
Speaking	In-class	124	32.94	20.86
	Online	124	9.15	12.08
Listening	In-class	124	28.98	19.51
	Online	124	25.16	27.30
Reading	In-class	124	19.36	16.44
	Online	124	27.35	24.33
Writing	In-class	124	18.52	18.09
	Online	124	38.49	25.18

*p <.05

(Adapted from De Veas et al., 2020)

Although the results comparing students' perceptions of the skills used in in-class and online lessons are not especially surprising, there are a few notable points of interest. As one might expect, reading and writing were perceived to be used much more in the

online lessons than in the in-class lessons and these differences were statistically significant; reading, $F(1, 123) = 9.30$, $p = 0.003$, and writing, $F(1, 123) = 77.04$, $p < 0.001$. The difference between perceptions of the use of speaking between online and in-class was also significant ($F(1, 123) = 123.29$, $p < 0.001$). However, if anything, the mean value of 9% speaking in the online lessons is unexpected given there was no spoken component in the online lessons. The only skill which was not statistically significant in its difference between online and in-class was listening, which was perceived to be used almost the same amount. This is also somewhat surprising given that the online lesson lacked spoken interaction and only used written instructions. Thus, these results may reflect the extensive use of video and audio texts in the Freshman English online lessons.

Students perceived the online lessons to be much less interactive than the in-class lessons both in terms of interactions with their peers (in-class: $M = 3.57$, $SD = 0.86$; online $M = 2.21$, $SD = 1.68$) and with their teacher (in-class lesson: $M = 1.86$, $SD = 1.35$; online: $M = 0.60$, $SD = 1.13$). These results were also both statistically significant (interactions with peers: $F(1, 23) = 71.35$, $p < 0.001$; interactions with teachers: $F(1, 123) = 83.10$, $p < 0.001$). Although it is unsurprising that there was less interaction in the online lesson, this may be important to students' ratings of the evaluative items below.

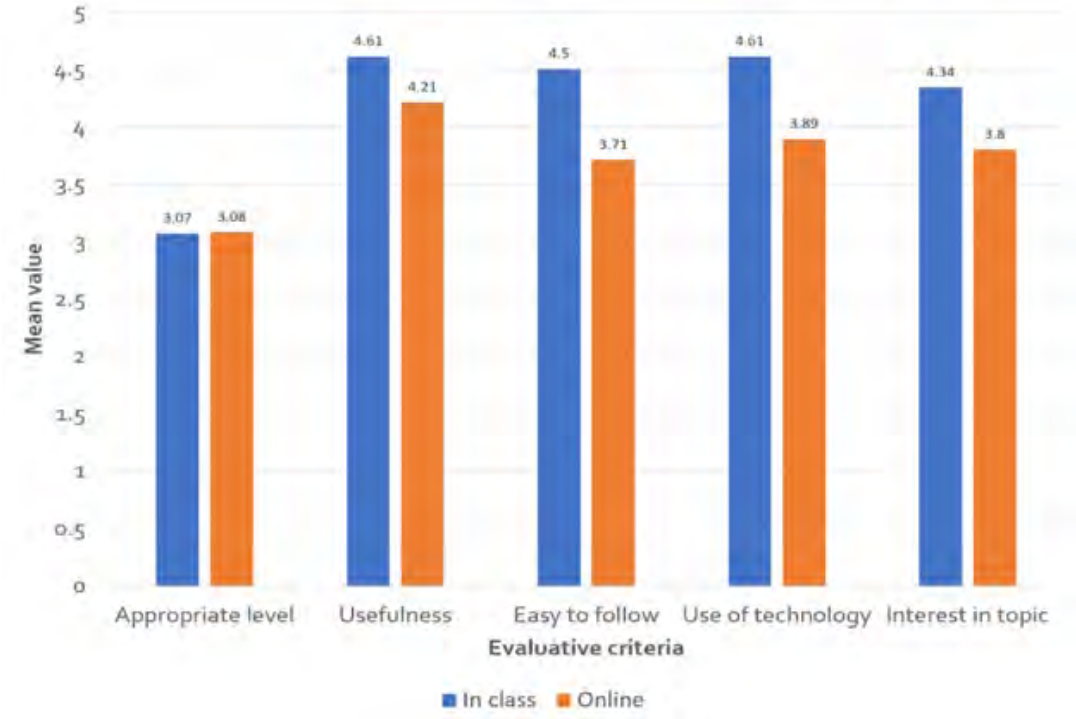
Student Evaluations

Student perceptions of the comparable value of in-class and online lessons were recorded using rating scales for level appropriateness, usefulness, ease to follow, use of technology, and interest. The mean values and standard deviations for each of the rating questions were recorded and an rANOVA was run to compare if the difference between in-class and online conditions was statistically significant. The difference was found to be statistically significant in students perceptions of usefulness ($F(1, 123) = 13.64$, $p < 0.001$), the ease to which they could follow the materials ($F(1, 119) = 32.65$, $p < 0.001$.), the use of technology ($F(1, 123) = 25.26$, $p < 0.001$), and the level of interest ($F(1, 123) = 21.18$, $p < 0.001$), with students perceiving the in-class lessons to offer greater value. There was no significant difference between student perceptions of level appropriateness.

Although, overall, the in-class lessons scored more highly on the evaluative criteria, it would be wrong to conclude that the online lessons were viewed negatively. In terms of the usefulness, ease to follow, use of technology, and interest in the topic, the mean values for the online lessons were positive, all rating higher than a three on each scale, which is above the midpoint of 2.5. Overall, therefore, students' perceptions of online lessons were still positive, even if they did not rate as highly as the traditional, in-class lessons.

Figure 1

Comparison of Mean Values of Evaluative Items.



(Reprinted from De Veas et al., 2020)

Qualitative Analysis of Student Perceptions of the Online versus In-person Lessons

Students' Comments

The wide variety of comments students provided in the survey mirrored the varying perceptions recorded by the evaluative items. These highlighted the difficulties and advantages of the online lessons. Many students stated that they enjoyed the online lesson finding it useful and interesting, especially in practicing their writing skills. However, several students mentioned that they found instructions difficult to follow online and some stated their preference for meeting in person, especially with reference to speaking. Overall, comments provided a selection of positive and negative views on the online format.

Table 2

Students' Comments from the Online Survey

Positive Student Comments:
Good lesson! So interesting
It is a good way to learn myself.
If I can see a example when I write something such as research paper, it is helpful.
The lesson has some advantages, but has also disadvantages I think.
The topic was really interesting and shocking for me.I was surprised the technology in this site.
いい経験になった [It was a good experience]
オンライン授業のメリットは、体調が悪くて家にいてもできるという点である。学校に出向くことができなくてもみんなと同じ授業を受けられるのは革新的だ [The good point about online lessons is that if you are unwell you can take them at home. It's innovative that we can take a lesson together without going to campus]
Negative Student Comments:
I was confused to do this class.
I want to take lesson from you in direct.
課題が時間内に終わらないので減らして欲しい [I couldn't complete the materials in time so I would like you to decrease them]
レッスンというよりただ動画を見せているだけでは?とも思う。ライティングの勉強にはなるがそれを添削してくれるわけではないし。[I think it was more like just watching videos. It wasn't like regular writing practice as you couldn't get corrections]
この授業スタイルなら家でもできる。ホームワークと変わらない非効率な授業に思えました。[We could complete this lesson style at home. It was no different to homework so I think it was an inefficient lesson]

Discussion and Limitations

Challenges for Teachers

Canvas is a vast, polished and professional LMS which is easy to use and intuitive for learners and teachers. The system allows teachers to create pages with multiple embedded media and then link these to discussion feeds easily, although, developing online lessons was found to be time consuming by the researchers. The lessons created were simple, clear and intuitive. However, it became apparent that lesson instructions needed to be very precise as it was not possible to deal with unanticipated misunderstandings given that students were away from the teacher and teachers were not able to check work until it was completed. The teachers also felt that they were unable to monitor students effectively, especially with respect to maintaining the university's English-only policy.

Using the free version of Canvas also limited which functions the researchers were able to include in their lessons. The paid version of Canvas may have enabled inclusion of other activity types. Over the last six months (since the start of emergency remote teaching) teachers' knowledge has improved and students have become more accustomed to online

learning. Given this experience, the addition of apps that allow for more student-to-student asynchronous interaction would be easier to integrate into the curriculum in the future, and potentially prove beneficial.

Challenges for Students

An additional area that needed extra care was in the initial student set-up procedure, as students had to first go into the Canvas settings and register their email addresses so that the teacher could then add them to the course and give them access to the materials. Doing this remotely was potentially time consuming and difficult for some students, especially in an environment in which they could not easily ask for help. As well as not being able to ask the teacher, students were also often missing the benefits of peer support as they completed the online materials. Completing this part of the set-up together in-class was preferable as a few students in most classes still had difficulty logging on and finding the materials despite using instructional videos and in-class directions.

This unfamiliar LMS, whilst allowing for a similar experience across all classes, added a burden for students already comfortable using other platforms previously during in-person classes. This may have also negatively impacted students' perceptions of the online component of the blended learning lessons. If this study were to be repeated it may be better to choose an LMS familiar to students and teachers to better compare the perceived value of online and in-person lessons.

Overall, however, students' comments on the survey included both positive and negative views on the online format. Also, in terms of the lesson design, it was positive that there was not a significant difference in the appropriateness of the level between the online and in-person lessons. This suggests that the lessons were well-designed for what the students were expected to be able to complete.

Conclusion

In this research project the comparative value to students and teachers of online and in-person lessons was evaluated using a mixed methods approach. Quantitative analysis of students' perceptions of the value of each lesson format showed that in-person lessons were rated significantly more valuable than online lessons in terms of usefulness, ease to follow, use of technology, and interest in the topic. This more positive evaluation may have been because in-person lessons were also considered to be more interactive by students. However, even though online lessons were not viewed as valuable as in-person lessons, they were still viewed positively overall. Students' comments also showed a range of positive and negative opinions about the online lessons which were reflected in the evaluative ratings. From a teacher's perspective, designing and creating the online lessons, although not difficult, was more time consuming than in-person lessons, and it was found that care especially had to be taken to give very precise and clear instructions. It should be noted that teachers needed to make sure that the goals of the lesson fitted with the affordances of the LMS, and that all students, including absentees, were well-orientated to the technology. If this was achieved, then teachers felt that effective and engaging materials could be created online despite these

lacking a spoken interactive component. Online lessons, as part of a blended learning environment may, therefore, be an option in providing better access to students for lessons in which in-person spoken interaction is not required.

References

Bollinger, A. (2017). Foreign language anxiety in traditional and distance learning foreign language classrooms. (Doctoral dissertation, Liberty University).

De Freitas, S.I., Morgan, J., & Gibson, D. (2015). Will MOOCs transform learning and teaching in higher education? Engagement and course retention in online learning provision. *British Journal of Educational Technology, 46*(3), 455–471.

De Veas, K., Frazier, E., Jurns, K., Asquith, S., Lege, R., & Lyon, P. (in press). In-class or online? Investigating student perceptions of asynchronous online lessons [Manuscript submitted for publication]. *Literacies and Language Education: Research and Practice 3*.

Garnham, C., & Kaleta, R. (2002). Introduction to hybrid courses. *Teaching with Technology Today, 8*(6). **http://www.uwsa.edu/ttt/articles/garnham.htm**

Garrison, D.R., & Kanuka, H. (2004). Blended learning: Uncovering its transformative potential in higher education. *Internet and Higher Education, 7*(2), 95–105.

Garrison, D.R., & Vaughan, N.D. (2011). *Blended learning in higher education: Framework, principles, and guidelines.* Jossey-Bass.

Huang, H. (2000). Instructional technologies facilitating on-line courses. *Educational Technology, 40*(4), 41–46.

Jeschofnig L., & Jeschofnig P. (2011). *Teaching lab science courses online: Resources for best practices, tools, and technology.* Jossey-Bass.

Johnson, N.H., Lyddon, P.A., Nelson, M.E., Selman, A., & Worth, A. (2015). JALT forum: Reimagining contemporary EFL curricula. In P. Clements, A. Krause, & H. Brown (Eds.), *JALT2014 conference proceedings.* JALT.

Kenney, J., & Newcombe, E. (2011). Adopting a blended learning approach: Challenges, encountered and lessons learned in an action research study. *Journal of Asynchronous Learning Networks, 15*(1), 45–57.

Kintu, M.J., Zhu, C., & Kagambe, E. (2017). Blended learning effectiveness: The relationship between student characteristics, design features and outcomes. *International Journal of Educational Technology in Higher Education 14,*(7). **http://doi.org/10.1186/s41239-017-0043-4**

Koszalka. T., & Ganesan, R. (2004). Designing online courses: A taxonomy to guide strategic use of features available in course management systems (CMS) in distance education, *Distance Education, 25*(2), 243–256. **https://doi.org/10.1080/01587910420000262111**

Mynard, J., & Murphy, P. (2012). An investigation into offering flexible language courses utilising blended learning. *Studies in Linguistics and Language Teaching, 23,* 101–137.

Twigg, C. A. (2003). Improving learning and reducing costs: New models for online learning. *EDUCAUSE Review, 38*(5), 29–38.

Wells, G. (1999). The zone of proximal development and its implications for learning and teaching. In *Dialogic inquiry: Towards a sociocultural practice and theory of education* (pp. 313–334). Cambridge University Press.

Authors' Bios

Steven Asquith is a senior lecturer at Kanda University of International Studies and co-editor of JALT's *The Language Teacher,* My Share column. He is interested in training learners to be more autonomous, especially relating to learning vocabulary, digital pedagogies, and new learning design.

Phoebe Lyon is a Principal Lecturer for Curriculum and Assessment at Kanda University of International Studies, Japan. Her research interests include learner autonomy, learner identity, materials development, and assessment.

Kathryn Jurns is a lecturer at Kanda University of International Studies. She received her Master of Education in Adult and International TESOL from the University of Cincinnati. Her research interests include academic writing in EFL, peer evaluation, and foreign language speaking anxiety.

https://doi.org/10.37546/JALTSIG.CALL2020.3

3

The Current State of MALL in Japanese Universities

Edo Forsythe, Hirosaki Gakuin University

Abstract

It has been almost 20 years since Thornton and Houser (2002) investigated the potential of using cellphones in Japanese university language classrooms. Since then, thousands of research articles have demonstrated the practical, theoretical, and pedagogical foundations of mobile-assisted language learning (MALL) in Japanese educational contexts. This chapter is a review of recent literature in the field of MALL, specifically focusing on the use of iPods, iPads, and smartphones in language learning. The research from the previous seven years reviewed herein includes studies about the use of video creation, social media, content instruction, and student perceptions of MALL. The intent of this chapter is to provide an overview of current MALL practices in use in Japanese educational contexts with ideas for research-based, practical application and to provide suggestions for continued research in this field.

ソーントンとハウザー（2002）が日本の大学語学の授業での携帯電話活用の可能性に関する研究を始めてから、もうすぐ20年になる。それ以後日本の教育環境におけるモバイル・テクノロジーを使った外国語教育方法（MALL）の実践的、理論的、教育的な基礎に関する研究論文が何千と発行されてきた。本章では最近のMALL研究、特にiPod、iPad、スマートフォンなどの活用に関する研究がまとめられている。7年前からビデオ制作やソーシャルメディア、内容重視教育、学習者及び指導者のMALLに関する認識調査等の研究論文が含まれている。本章の目的は、日本の教育環境で活用されているMALLの実践の概要を提供することと、継続的な研究を促すことである。

Keywords: **MALL; CALL; mobile language learning; video; SMS; iPad; smartphone; Extensive Reading**

It has been almost 20 years since Thornton and Houser (2002) investigated the potential of using cellphones in Japanese university language classrooms. Since those days, thousands of research articles have demonstrated the practical, theoretical, and pedagogical foundations of mobile-assisted language learning (MALL) in Japanese university contexts. Today's mobile technology has brought a revolution in access to world cultures and languages for language learners; consequently, language teachers are experimenting constantly to make the best use of such technology in the language classroom. Specifically, smartphones have become a common tool in Japanese university EFL classrooms, and many teachers are requiring students to use them to perform language learning activities (Barrs, 2011; Davies,

2015; Garcia Mendoza, 2014; Lockley, 2013; Nakaya & Murota, 2013; Stockwell & Liu, 2015). As Oberg and Daniels (2013) noted, the majority of research into the use of mobile phones in language learning conducted prior to their study was conducted using feature phones – the type of mobile phones which preceded smartphones such as the iPhone, Android, Galaxy, and others that have a touch screen, an interactive operating system, and high-speed, wireless internet connectivity (Wu, 2015). Cote et al. (2014) found that almost 100% of Japanese university students now own personal smartphones, providing an opportunity to use these devices in university classrooms. Along with this increase in student smartphone ownership, Burston (2013) posited that smartphones have brought a new age of MALL with a variety of possibilities. This brief review of recent literature will explore the current state of MALL focusing on the state of smartphone use in Japanese university EFL education, and student perceptions of MALL methods being used in their EFL classrooms.

Documentation

The various research papers which support this study were found through multiple means. Online databases were searched using the following terms to find relevant materials and articles: mobile technology, mobile device, mobile phone, smartphone, cellphone, language learning, English, EFL, foreign language, MALL, CALL, and technology. In addition to online research database searches, regular reviews of the archives of pertinent journals were also conducted because their contents were not always indexed in the databases. Examples of these publications include the *JALT Journal*, *The JALT CALL Journal*, *The Language Teacher*, and other refereed academic publications associated with foreign language education in Japan. It must be noted that this review of existing research is limited to research published in English and does not include the wealth of information published in Japanese about MALL in Japanese university contexts.

Mobile Devices in Japanese EFL Studies

With the development of mobile technology such as tablet computers, iPads, cellphones, smartphones, and portable audio devices, instructors using CALL teaching methods quickly began to experiment with the inclusion of the new mobile technology in the language classrooms (Burston, 2013). Many educators at Japanese universities have been exploring the impact that mobile devices might have on their EFL courses (Brown et al., 2012; Paterson, 2014; see also Ockert, 2013; Ockert 2014 for use of iPads in high schools). These studies will be explored in detail below, starting with the use of iPods.

iPods in Japanese University EFL Classrooms

In an experimental, quantitative study using iPods – mobile devices which offer nearly identical learning opportunities as smartphones – in Japanese university EFL classes, Oberg and Daniels (2013) explored the degree to which self-paced learning using university-issued mobile devices affected students' learning. They found that the experimental groups, which were permitted to study at their own pace in the classrooms, performed

better on the standardized assessments than the control group participants. Another study which investigated the effect of an iPod-based language learning application on Japanese university students' EFL studies was conducted by Cortez and Roy (2012). Their qualitative case study investigated how the interface of a mobile application affected students' success in MALL activities. Through focus group interviews, Cortez and Roy found that Japanese university students were comfortable and successful in completing EFL learning tasks using mobile devices. However, poorly designed user interfaces were revealed to inhibit their participants' success in learning. These iPod-based MALL studies lend further credence to the above-mentioned findings that mobile devices can have a positive effect on EFL learning in Japanese universities.

iPads in Japanese EFL Education

Brown et al. (2012) reported on one of the first studies using iPads in Japanese university EFL courses after the introduction of the Apple iPad in 2010. Their qualitative study explored how students felt about the introduction of iPads into their EFL classes for a variety of classroom tasks, some of which had previously been completed on computers. As is often the case of introducing new technology into a process, Brown et al. found that students encountered some problems with using the iPads. Specifically, the participants reported that they saw no need to use technology devices in activities which were easily accomplished using a pencil and paper – a notion echoed later by Pegrum (2015) that technology should not be introduced for technology's sake. Demonstrating that careful consideration of how and whether the technology will improve the educational process is necessary. Having found that students and teachers alike enjoyed using iPads in some EFL activities and seeing the potential for positive impact on English education in Japanese universities, Brown et al. laid the groundwork for future research into effective iPad integration. In addition to the high school-based studies by Ockert (2013; 2014) mentioned above, another such study was conducted by Paterson (2014) in a private university's EFL program.

Paterson (2014) was interested in the plausibility of students using a set of iPads in academic EFL writing and reading classes in an attempt to confirm previous research in the field that was consistent with personal experiences. In a qualitative approach to analyzing survey data collected at the end of a semester of iPad use, Paterson found that the majority of the participants thought that the iPads were useful in their English studies and saw the potential for future efficacy of the devices, thereby confirming Brown et al.'s (2012) and Ockert's (2013) research. Specifically, Paterson's (2014) results revealed that the participants particularly enjoyed the mobile aspect of the iPad, in that they could use them while traveling or waiting for transportation – anytime they had free time in non-traditional study situations. It is precisely this mobile, anywhere-learning aspect of MALL that appeals to both students and educators; and because these characteristics are also found in smartphones, teachers have begun to take advantage of the fact that almost 100% of Japanese university students own smartphones (Cote et al., 2014; Lockley, 2011; 2013; Murray & Blyth, 2011).

Smartphones in Japanese EFL Learning

As new mobile devices are released, teachers of Japanese EFL courses have taken their increased capabilities into consideration when developing language learning activities, so that the activities have gone from simple vocabulary study activities on feature phones – non-smartphones – (Stockwell, 2007) to video creation (Barrs, 2011) and interactive reading and grammar practice activities on smartphones (Wang et al., 2016; Wang & Smith, 2013). The section below explores recent research into the use of smartphones in Japanese university EFL courses.

Students having smartphones has become the norm in Japanese university EFL classrooms, so the teachers are taking advantage of these powerful resources. Accordingly, research into the use of smartphones in EFL classes has grown as well. In a qualitative survey of the effects of encouraging language learners to use smartphones, Barrs (2011) found that Japanese college students were using smartphones as a tool to improve their learning strategies, such as notetaking, lecture recording, and EFL vocabulary studying.

As for research into specific activities that Japanese university EFL students engage in using their mobile devices, Stockwell (2010) investigated how students used their mobile phones – feature phones in this case – to interactively study English vocabulary. In this quantitative study, Stockwell found that there were no significant differences in language learning performance between vocabulary learning activities conducted on a mobile device versus those performed on a computer, in spite of those performed on a mobile device taking more time to complete. One important revelation of Stockwell's research was that over similar research conducted in previous years, slightly more students were willing to attempt vocabulary learning activities using their mobile devices. Continuing this line of research in a smartphone-focused, mixed-methods comparison of Japanese and Taiwanese university students, Stockwell and Liu (2015) learned that Japanese university students are increasingly willing to use their smartphones in EFL learning activities, and that smartphones allow users to perform comparably to performing on computers on language learning activities.

Some of the changes in mobile learning habits observed by Stockwell and Liu (2015) confirmed the findings of Wang and Smith (2013) in their mixed-methods study of Japanese university students' study of EFL reading and grammar using smartphones. Two-thirds of their participants reported that they enjoyed the EFL reading activities performed on their smartphones, so long as the content was appropriate to the participants' interests and ability level. Reading and grammar materials viewed as too demanding were not as successfully accomplished by the participants, showing that great care must be used when developing materials for mobile device-based study. Overall, Wang and Smith found that a wide majority of Japanese university students find MALL-based EFL learning activities helpful in improving their foreign language abilities.

Yamauchi and Uchida's (2011) exploration of existing research into the use of smartphones in the EFL classroom reported that using smartphones in Japanese university EFL courses helps students overcome their lack of computer skills and that students preferred to perform communicative activities over language drills when using their smartphones. Leis et al., (2015), in a mixed-methods, non-experimental study of how smartphones influence

learner autonomy, supported using smartphones in language learning, finding that using smartphones in class encourages more language study outside of the classroom.

This long history of the use of smartphones in Japanese university EFL education has been met with a generally positive response as university students slowly adapted to using their devices for educational purposes in addition to their main function of personal communication. The research of Wang et al. (2016) follows the recent history of MALL in Japanese university English education through their Mobile English Learning Project – a program in which EFL study materials were delivered through students' mobile phones. In their quantitative research into their participants' perceptions and preferences for learning via mobile devices, Wang et al. found that Japanese university students in general welcome the idea of EFL study via their mobile phones and feel that it is beneficial (p. 436). However, this study also revealed that students will abandon voluntary EFL studies when their schedules become full, so teachers who hope to engage students in MALL activities should integrate them into their curricula in order to maintain student engagement (Wang et al., 2016).

These findings follow with those of Daniels (2012) in a quantitative survey of Japanese university students' mobile devices and their usage patterns. Daniels reported that, like Wang et al. (2016), a large majority of Japanese university students are interested in using their smartphones for EFL studies, and Daniels posited that their devices can be used in place of traditional language laboratories where students used technology to study or practice their foreign language skills. The interactional aspects of smartphones have been especially used in MALL to take advantage of students' communications practices in social media as explored below.

Japanese University Students' Use of SMS, SNS, and Web-based Platforms

In order to address a gap in literature about the effectiveness of short message service (SMS) and social networks (SNS) for communication via mobile phones, Garcia Mendoza (2014) explored whether mobile phones enabled increased interaction in online discussions and how interacting via mobile phones differed from desktop computers. Garcia Mendoza found that the participants in both the mobile device group and the desktop group shared similar content, but the students' interactions were more common on mobile devices but the messages were of shorter lengths (p. 229). Garcia Mendoza suggested that this may be due to the immediacy of discussions via mobile devices, whereas desktops are more conducive to longer, well-articulated responses to comments.

Similarly, Wu and Marek (2016) investigated the use of the LINE SMS application and email for student collaborative writing activities between Japanese and Taiwanese university students. Their mixed-methods study found that students successfully used LINE to communicate with other group members and email to coordinate their writing assignment progress. Wu and Marek's research revealed that the SMS interactions increased the students' motivation to communicate and collaborate cross-culturally (p. 62). The results of analysis of the participants' experiences by Garcia Mendoza (2014) and Wu and Marek (2016) were similar to those found by Daniels (2012), Gromik (2012), Nishio and Nakatsugawa (2020), and Wang et al. (2016), that mobile phones have great potential to enhance interaction in online collaboration.

Other Studies Into SNS Use in EFL Contexts

At the same time as Garcia Mendoza (2014), Leis (2014) examined student perceptions of using Twitter in an EFL class for both consuming content and communicating through their own tweets and comments. This quantitative study found that the incorporation of such social media in the EFL classroom had a positive effect on student's autonomy and motivation to use SMS in English (p. 75). Researching another popular SMS platform – Facebook – Ohashi (2016) studied how using Facebook in EFL classes could provide increased opportunities to communicate in English, provide English-language resources, motivate students to use English, and create a learning community for them (pp. 346–348). The findings indicated that positive results in all four categories of study were found, and therefore, supported Leis' finding that using SMS in EFL courses has many benefits for English learners.

The research conducted by Garcia Mendoza (2014), Leis (2014), and Ohashi (2016) made a compelling case for the use of mobile devices for student collaboration and interactions on social media and demonstrated that using smartphones in MALL-based activities was beneficial for the students. However, communications and interactions are not the only MALL tasks which can be performed on smartphones.

Video in MALL

In addition to MALL methods being used for online intercultural collaborations via the internet, instructors are also using technology to facilitate the creation and sharing of student videos in their foreign language education. From Gromik's (2012) and Hirotani and Lyddon's (2013) research into video self-introductions to Hirschel et al.'s (2012) video self-assessments, to Ockert's (2014) and Toland et al.'s (2016) use of videos as an intervention tool in presentation preparation, and relatedly, Hawking's (2014) use of smartphones to record and submit audio files, educators have taken advantage of the ability of students to use the audio-visual medium in their language learning activities.

Gromik (2012) used a qualitative, single case study approach to understand how Japanese university students felt about using cellphones to create videos of themselves speaking English. This small-scale study demonstrated that the students enjoyed making videos and that the activities increased their motivation to excel in their English studies. Similarly, in their mixed-methods study of the influence of doing video self-introductions in an asynchronous online intercultural collaboration between American and Japanese university students, Hirotani and Lyddon (2013) found that creating the videos had a positive effect on the language development of their participants. Hirotani and Lyddon found that sharing videos about oneself and one's culture enabled students to gain both linguistic and cultural knowledge. Further, the post-study survey revealed that the participants enjoyed making and watching the videos in spite of minor technical issues often associated with MALL activities.

Hirschel et al. (2012) also found that Japanese university students enjoy using videos in their EFL courses in their quantitative study of the effect of video self-assessments on EFL language learners' language ability. Building on the idea that self-assessment can have

significant beneficial effects on language learners' development, Hirschel et al. (2012) explored the use of self-video in students evaluating their own foreign language abilities. The study revealed that after having created and watched several videos of themselves speaking English, the students experienced gains in their interest in English and in their self-confidence in using English, in spite of minimal development in linguistic or lexical fluency (p. 304). This study demonstrated the benefit of using videos in the EFL classroom, an activity that is even easier to do with mobile devices, as Ockert (2014) also found in a longitudinal study of using iPads in the EFL classroom for a variety of activities, including video creation. In a quasi-experimental, quantitative study, Ockert learned that using iPads to video record Japanese high school student presentations led to increased confidence in English ability and willingness to communicate in English, as well as a decrease in anxiety when compared with recordings made using a traditional camcorder (p. 49). The ease in playback of the videos was reported as a potential reason for the decreased anxiety and increased affect, because the participants could immediately critically view themselves speaking and assess their own development (Ockert, 2014, p. 63). In addition to Ockert's 2014 research, a previous study into the use of iPads found similar positive effects on EFL students' learning (Ockert, 2013).

Toland et al. (2016) also examined the use of videos recorded using mobile devices for enhancing English students' presentation skills by having them conduct self-reflection and peer review activities. Their quantitative study demonstrated that students found the use of mobile video to be beneficial in improving their presentation skills and their overall English abilities. The benefits of the video-creation capabilities of mobile devices have been clearly found by the research of Gromik (2012), Hirotani and Lyddon (2013), Hirschel et al.'s (2012), Ockert's (2013; 2014), and Toland et al.'s (2016). These examples provide rich ideas for educators to follow their lead by having students create videos in their EFL study. However, mobile devices are not only for language practice, they can effectively be used in content instruction as well, as discussed in the following section.

Content-based EFL Instruction Via Smartphones

Daniels (2012) demonstrated that smartphones can be used by Japanese university students to access course content hosted on learning management systems like Moodle and Blackboard. The network connectivity options of smartphones and the prevalence of unlimited data plans among Japanese university students (Wang et al., 2016) has enabled the students to use their devices as learning tools. Ko et al. (2015), in their multinational quantitative survey of university students, found that the students often use their devices in information-gathering activities and for submitting course assignments in addition to the aforementioned collaboration activities. Ilic (2015) also identified that students spent more time on their EFL-related tasks because they were completing them on their smartphones. Although Ko et al.'s (2015) participants reported little experience in using their smartphones for academic reading, Milliner (2015) reported success in the students' performing extensive reading tasks on their smartphones.

In his quantitative study of the effects of mobile device-based extensive reading (ER) tasks on Japanese university students' EFL proficiency, Milliner (2015) revealed that taking

advantage of the proliferation of mobile devices among the students can broaden their access to ER language study materials. Despite a number of limiting external factors, Milliner was able to state that the participants' overall reading speed increased after participating in a ER-focused course that relied on the students' mobile devices – smartphones and tablets – to deliver the content. Further, an increase was noted in the students' reading scores on the Test of English for International Communication (TOEIC) after completion of this course (naturally, Milliner did not draw a causal relationship to the MALL activities, but it could be hypothesized that the MALL methods were not detrimental to the students' general proficiency). In a separate but related quantitative study, Milliner (2017) found that Japanese university students preferred using their smartphones for ER assignments over using a personal computer at a rate of 97% (p. 54). The research of Daniels (2012), Gromik (2012), Ko et al. (2015), Milliner (2015; 2017), and Wang et al. (2016) has demonstrated that MALL using smartphones in Japanese university EFL education is not limited to communications-type activities, but includes rich options for content creation, research, video recording, and even extensive reading. Lee (2019b) even explored the use of smartphones for EFL writing activities, finding that students produced less text when writing on smartphones than on computers, even though they enjoyed using their devices for English study (p. 229). With these activities being incorporated more often in EFL classes, it is important to understand how the students feel about this methodological shift and the use of their personal devices.

Student Perspectives on Mobile Learning and Social Media

MALL methods have been generally accepted and enjoyed by Japanese university students for years, with stronger students reported by Kobayashi and Little (2011) as benefitting more from the blended learning methods. Burston (2013) saw this use of student devices as the inevitable future of MALL in Japanese university EFL classes. With the obsolescence of traditional CALL classrooms and language laboratories, and with the continuously improving ability of apps and programs to work across different operating systems, teachers are turning to student smartphones to continue the blended learning instruction. As Burston's (2013) future has come to pass, it is important to consider how students feel about such demands on their personal devices.

Gikas and Grant (2013) explored the perceptions of the students in using their own mobile devices, as well as the ways in which mobile devices were affecting learning at universities. The purpose of this qualitative case study was to better understand how educational situations were affected when mobile devices were implemented in university classrooms. This study presented research findings regarding students' perceptions of learning with mobile devices and the roles social media played in the educational process. Gikas and Grant's qualitative, non-experimental research focused on American university students who had experience using mobile devices into their language courses. The implications for the field of mobile educational technology are that mobile devices and social media can be leveraged to encourage interaction, collaboration, and content creation among students. Gikas and Grant's (2013) research provided further evidence of the positive effects that mobile technology has on the learning process and laid out specific examples of why teachers should learn

how to include mobile technology into their lessons. The research methods made a compelling case for the significantly positive impact mobile technology has had on language learning, setting the stage for similar research to be conducted in Japanese universities.

Student Perceptions of Smartphones in Japanese EFL Classes

In a study of Japanese university EFL students' use of cellphones to study vocabulary, Stockwell (2010) found that the participants were capable of using mobile devices to learn, but that perhaps they were not ready to use mobile phones in their EFL education because many of the participants opted to use their PCs instead of their mobile devices to complete language-learning tasks. However, with the advent of the smartphone, this attitude has been changing and Japanese university EFL students have begun to embrace MALL methodology (Wang & Smith, 2013; White & Mills, 2014; Lee, 2019a).

Barrs (2011) demonstrated that Japanese university students are using their personal smartphones as language learning tools, which supported Stockwell's (2007) statement of the students' ability to learn via cellphone. Quantitative studies by White and Mills (2012, 2014) surveying Japanese college student attitudes toward EFL learning via smartphone reported that student perceptions of using smartphones in MALL continue to improve. Gromik (2012) revealed in a single, qualitatively analyzed case study of Japanese university students that the participants understood the benefits of using mobile phones in language learning and that they were good learning tools. In a mixed methods survey of Japanese university students, Lockley (2013) also found that the participants perceived the value of using mobile technology in the EFL classroom and that students felt more familiar with smartphone technology than with personal computers.

In a similar study that quantitatively explored Japanese university students' attitudes toward CALL methodology, Lockley and Promnitz-Hayashi (2012) revealed that the participants expressed a desire for MALL activities in their EFL classes. Further, in a case study of Japanese university students' technology preferences, Garcia Mendoza (2014) reported that participants saw smartphones as a better tool for collaboration than personal computers, but that computers were preferable for academic work.

While most of Wang and Smith's (2013) participants indicated that they had never used their mobile phones to learn anything before, they did report an enjoyment and benefit of performing grammar and reading activities on their smartphones. However, in interviews, Wang and Smith's students stated that they had difficulty reading and performing assessments on smartphone screens, and expressed an opinion that, in addition to privacy and security concerns, academic study should be done in class or on a PC and that smartphones were more for personal activities. So, there is some question as to whether Japanese university students feel comfortable with teachers' requirements or expectations that students use their personal devices to perform language-learning activities.

Along the lines of Wang and Smith's (2013) findings that students had some hesitations about using smartphones in EFL study, Lee (2019a) found in a quantitative study of Japanese university students at a technical university that the majority of the students were not opposed to using their smartphones, but that a significant number of them were reticent to adopt some smartphone-based EFL activities due to their personal preferences

for more traditional learning methods. This attitude was especially found in writing tasks – numerous students preferred to write by hand instead of typing into their smartphones (also reported in Lee, 2019b).

Wang and Smith's (2013) findings were supported by Ko et al. (2015) in a quantitative survey of four countries' (Japan, Taiwan, Hong Kong, and America) university students regarding their capabilities in using technology. Contrary to what Wang and Smith (2013) learned from their study, Ko et al. found that the participants were capable of using mobile devices for collaboration but not for effectively performing reading or academic work. From a mixed methods study of Japanese university students' mobile collaboration, Ilic (2015) also reported that smartphones facilitated collaboration among group members in EFL classes, and that students were positive about using smartphones for on-the-go studying. Thus, it could reasonably be posited that the most effective use of smartphones in the Japanese university EFL classroom is for collaboration more than for actual studying.

One concern that language learning on smartphones may not be as effective as other technologies is the intersection of students' personal and academic lives on their smartphones. Because students use their devices for personal communications as well as for information retrieval and task accomplishment, Hill (2015) explored the potential for the personal communications to become a distraction to students' learning. In a quantitative study of Japanese university students studying in EFL classes, Hill found that a majority of the study's participants thought that they were able to responsibly use their smartphones in the language classroom without being distracted: they felt little inclination to play games during class, nor were they overly compelled to interrupt their learning to respond to an SMS message during class (p. 27), thereby dispelling the concern. Further, the use of smartphones by some students in the EFL classroom proved not to be a significant distraction to other students in the class, as opposed to the reporting of Sana et al., (2013) which stated that laptops in the university classroom can be a distraction to other students.

The matter of ownership of the smartphones may be relevant in the area of this MALL in Japanese university EFL education. Oberg and Daniels (2013) and Cortez and Roy (2012) found in their quantitative studies using iPods issued to the students by the university, that the participants' attitudes toward MALL were very positive. It is possible that the fact that students used technology that was issued to them changed their perception of the MALL activity because it did not impose on their personal devices and in their personal space. Wang and Smith's (2013) research revealed students' belief that mobile phones were personal and separate from their concept of classroom learning tools. Additionally, in a quasi-experimental, quantitative study of the use of mobile devices in EFL studies, Kondo et al. (2012) found that participants raised privacy concerns when asked to use their own smartphones, so the study was conducted with mobile gaming hardware instead. Wang et al. (2016) and Yang and Wang (2014) reported similar privacy concerns among their participants regarding the use of their personal smartphones and their associated personal information – email addresses, phone numbers, etc. – to access linked websites and register for online language learning materials. Privacy and ownership perspective are just two of the concerns that should be considered when asking students to use their personal smartphones for language learning activities.

Conclusion

The history of CALL methods being used in Japanese university EFL classrooms is almost as long as that of CALL in western universities. Thornton and Houser (2002, 2005) conducted some of the first research into the benefits of CALL; and naturally, as technology progressed and developed, so did its use in the EFL classroom. Cellphones and mobile devices have been used in Japanese university EFL classes since their advent, and teachers have studied how best to employ them in language learning activities (Stockwell, 2007; 2008; 2010).

Blended learning methods are quite common in Japanese university EFL education, and their use continues to grow (Hinkelman & Gruba, 2012; Kobayashi & Little, 2011). The use of MALL practices has had a generally positive effect on student motivation in the foreign language classroom (Ushioda, 2013), and has spurred autonomous learning, as demonstrated by Byrne and Diem's (2014) study of the demographics of language learning app users. The improvement of internet technology has enabled EFL students to participate in online intercultural collaborations (Carney, 2006; Forsythe, 2014; Flowers, 2015; Flowers & Kelsen, 2016), and web-based language-learning video creation activities (Gromik, 2012; Hirotani & Lyddon, 2013; Hirschel et al., 2012), and many other technology-based language learning activities. Such collaborations are now possible on the mobile devices that almost every Japanese university student carries with them – their smartphone.

Burston (2013) predicted that MALL is the future of language learning technology. Pegrum (2015) demonstrated that MALL can transform instruction into more student-centered learning. The use of mobile devices in Japanese EFL classrooms has been wide-spread and very effective through employment of a variety of devices. iPods were found to be efficient learning tools in the EFL classroom (Cortez & Roy, 2012; Oberg & Daniels, 2013); as were iPads in different settings (Brown et al., 2012; Ockert, 2014; Paterson, 2014). The effects of these devices were so positive that the researchers all highly recommended their implementation.

As cellphone and eventually, smartphone technology proliferated among the university student population, many instructors sought to take advantage of these new tools. Stockwell (2007; 2008; 2010) was an early adopter and researcher into the effective use of mobile phones in EFL education, finding that students could use their cellphones, but preferred computers for a variety of reasons. As smartphones became more common, researchers investigated the effectiveness of their use as well. Barrs (2011) found that smartphones could be used by students for notetaking and similar logistical support activities in their learning efforts. Lee (2019a; 2019b), Stockwell and Liu (2015), Wang and Smith (2013), Milliner (2015), as well as Wang et al. (2016) found that Japanese university students could use their smartphones for studying EFL grammar, vocabulary, and for practicing reading comprehension skills, and basic writing activities. Ko et al. (2015) and Ilic (2015) revealed the students' ability to learn content using their smartphones, but a large amount of existing research into MALL using smartphones involves studies of the devices being used for collaboration and communications. Gikas and Grant (2013), Garcia Mendoza (2014), Leis (2014), and Ohashi (2016) found positive effects of using social media in EFL studies, while

Ilic (2015), Ko et al. (2015), Nishio and Nakatsugawa (2020), and Wu and Marek (2016) revealed that smartphones are effective tools in collaborative learning in an EFL context.

Many of the studies discussed above have demonstrated that Japanese university students view the use of smartphones and mobile devices for EFL language learning positively and understand the benefits of using them (Gromik, 2012; Ko et al., 2015; Lockley, 2013; Lockley & Promnitz-Hayashi, 2012; White & Mills, 2012; 2014). This was borne out by the author's doctoral dissertation research which explored student perceptions of using their smartphones in EFL classes (Forsythe, 2017). Lockley and Promnitz-Hayashi (2012) also found that students want teachers to implement MALL, and White and Mills (2012, 2014) showed that students' desires for the future trends of Japanese university EFL programs include MALL methodologies.

Suggestions for Future Research

As demonstrated above, smartphones are prolific among Japanese university students and their presence in the EFL classroom is being taken advantage of by both the students and the teachers alike. This practice has become normalized to the point that students enjoy and even desire MALL activities in their language classrooms because they see the benefits of such practices. However, some studies have revealed some hesitation or disapproval among Japanese university students of being required to use their personal smartphones (Wang & Smith, 2013; Yang & Wang, 2014), further exploration of how students feel about using their smartphones in language classrooms is called for. Also, Hall (2016), Ilic (2015), Stockwell and Liu (2015), and White and Mills (2012, 2014) all recommended additional research into what student perceptions truly are of using their personal smartphones for EFL activities.

Finally, in addition to student perceptions of using smartphones in Japanese university EFL classrooms, teacher perceptions need to be explored as well. In preparing for this study, an extensive exploration of the literature in the field of educators' perceptions of mobile technology use in Japanese university foreign language learning programs revealed no published studies. Because, Hinkelman and Gruba (2012) found that the use of technology in the language learning classroom changes the dynamic of the classroom, including how the teachers teach and how and what students learn, it is important to consider the teachers' perceptions of how MALL could affect the EFL classroom environment. These types of additional research will add valuable information to the field of mobile assisted language learning in a Japanese university context and should be continuously explored so long as language teachers are using technology in their language classes.

References

Barrs, K. (2011). Mobility in learning: The feasibility of encouraging language learning on smartphones. *Studies in Self-Access Learning Journal, 2*(3), 228–233. **http://sisaljournal.org/archives/sep11/barrs**

Brown, M., Castellano, J., Hughes, E., & Worth, A. (2012). Integration of iPads into a Japanese university English language curriculum. *The JALT CALL Journal, 8*(3), 197–209. **http://doi.org/10.29140/jaltcall.v8n3.141**

Burston, J. (2013). MALL: Future Directions for BYOD Applications. *The IALT Journal, 43*(2). 89–96. **http://www.iallt.org/sites/default/files/07_techcolumn_iallt_journal_43-2.pdf**

Byrne, J., & Diem, R. (2014). Profiling mobile English language learners. *The JALT CALL Journal, 10*(1), 3–19. **http://doi.org/10.29140/jaltcall.v10n1.165**

Carney, N. (2006). Telecollaboration for intercultural learning: An overview of projects involving Japan. *The JALT CALL Journal, 2*(1), 37–52. **http://doi.org/10.29140/jaltcall.v2n1.21**

Cortez, R., & Roy, D. (2012). Screen interface design for mobile-assisted language learning in EFL context: A case study in Japan. *Journal of Language Teaching and Research*, *3*(3), 353–368. **http://doi.org/10.4304/jltr.3.3.353–368**

Cote, T., Milliner, B., Flowers, S., & Ferreira, D. (2014). What's going on at the MALL? *Peer-Spectives 12*, 37–40.

Daniels, P. (2012). Technically speaking: Moodle, mobile apps and mobile content. *The JALT CALL Journal, 8*(1), 33–44.

Davies, R. (2015). Why tweet when you can bubble? Students' perceptions of a voice microblog for the development of their L2 spoken production skills. *The JALT CALL Journal, 11*(3), 213–233.

Flowers, S. (2015). Developing intercultural communication in an ELF program through digital pen pal exchange. *The Center for ELF Journal, 1*(1), 25–39.

Flowers, S.D., & Kelsen, B.A. (2016). Digital sojourn: Empowering learners through English as a lingua franca. In P. Clements, A. Krause, & H. Brown (Eds.), *Focus on the learner*. JALT.

Forsythe, E. (2014). Online intercultural collaborations using wikis: An analysis of students' comments and factors affecting project success. *The JALT CALL Journal, 10*(3), 255–271. **http://doi.org/10.29140/jaltcall.v10n3.179**

Forsythe, E. (2017). A qualitative case study of Japanese university students and personal smartphone use in English as a foreign language classes (Doctoral dissertation). Available from ProQuest Dissertations and Theses Global database. (ProQuest document ID No. 2006913606).

Garcia Mendoza, G.A. (2014). A comparative study of computer and mobile phone-mediated collaboration: The case of university students in Japan. Mobile Learning Applications in Higher Education [Special Section]. *Revista de Universidad y Sociedad del Conocimiento (RUSC).11*(1), 222–237. **http://doi.org/10.7238/rusc.v11i1.1898**

Gikas, J., & Grant, M. (2013). Mobile computing devices in higher education: Student perspectives on learning with cellphones, smartphones & social media. *Internet and Higher Education, 19*. 18–26. **doi: http://doi.org/10.1016/j.iheduc.2013.06.002**

Gromik, N.A. (2012). Cell phone video recording feature as a language learning tool: A case study. *Computers & Education*, *58*(2012), 223–230. **http://doi.org/10.1016/j.compedu.2011.06.013**

Hall, J. (2016). Student perceptions of smartphone use for EFL learning. In M. Iguchi, & L. Yoffe (Eds.), *JACET summer seminar proceedings: Vol. 14. Mobile learning in and out of the classroom: Balancing blended language learner training* (pp. 24–30).

Hawking, R. (2014, June 6–8). *Recording and submitting audio recordings using mobile phones* [Conference Session]. JALTCALL2014 Conference, Sugiyama Jogakuen University, Nagoya, Japan.

Hinkelman, D., & Gruba, P. (2012). Power within blended language learning programs in Japan. *Language Learning & Technology*, *16*(2), 46–64. **http://llt.msu.edu/issues/june2012/hinklemangruba.pdf**

Hirotani, M., & Lyddon, P. (2013). The development of L2 Japanese self-introductions in an asynchronous computer-mediated language exchange. *Foreign Language Annals, 46*(3), 469–490.

Hirschel, R., Yamamoto, C., & Lee, P. (2012). Video self-assessment for language learners. *Studies in Self-Access Learning Journal, 3*(3), 291–309. **http://sisaljournal.org/archives/sep12/hirschel_yamamoto_lee**

Ilic, P. (2015). The effects of mobile collaborative activities in a second language course. *International Journal of Mobile and Blended Learning*, *7*(4), 16–37. **http://doi.org/10.4018/IJMBL.2015100102**

Ko, E., Chiu, D., Lo, P., & Ho, K. (2015). Comparative study on m-Learning usage among LIS students from Hong Kong, Japan and Taiwan. *The Journal of Academic Librarianship*, *41*(5), 567–577. **http://doi.org/10.1016/j.acalib.2015.07.005**

Kobayashi, K., & Little, A. (2011). Learner perceptions on the usefulness of a blended learning EFL program. *The JALT CALL Journal, 7*(1), 103–117. **http://doi.org/10.29140/jaltcall.v7n1.111**

Kondo, M., Ishikawa, Y., Smith, C., Sakamoto, K., Shimomura, H., & Wada, N. (2012). Mobile assisted language learning in university EFL courses in Japan: Developing attitudes and skills for self-regulated learning. *ReCALL*, *24*(2), 169–187.

Lee, B. (2019a). Japanese tertiary students' access to smartphones and their feelings regarding their use in the EFL classroom. *Memoirs of Fukui University of Technology, 49*, 216–224. **https://bradfordjlee.files.wordpress.com/2020/05/lee2019a_jp-tertiary-ss-access-to-smartphones-and-their-feelings.pdf**

Lee, B. (2019b). A case study of writing task performance: Smartphone input vs. handwriting. *Memoirs of Fukui University of Technology, 49*, 225–231. **https://bradfordjlee.files.wordpress.com/2020/05/lee2019b_a-case-study-of-writing-task-performance-smartphone-vs.-handwriting.pdf**

Leis, A. (2014). Encouraging autonomy through the use of a social networking system. *The JALT CALL Journal, 10*(1), 69–80. **http://doi.org/10.29140/jaltcall.v10n1.168**.

Leis, A., Tohei, A., & Cooke, S. (2015). Smartphone assisted language learning and autonomy. *International Journal of Computer-Assisted Language Learning and Teaching, 5*(3), 75–88. **http://doi.org/10.4018/IJCALLT.2015070105**

Lockley, T. (2011). Japanese students' experience of ICT and other technology prior to university: A study. *The JALT CALL Journal, 7*(1), 93–102. **http://doi.org/10.29140/jaltcall.v7n1.110**

Lockley, T. (2013). Answers to outstanding questions about Japanese student ICT competencies and a glance into a mobile future. *Asia-Pacific Education Researcher, 22*(4), 603–617. **http://doi.org/10.1007/s40299-013-0063-3**

Lockley, T., & Promnitz-Hayashi, L. (2012). Japanese university students' CALL attitudes, aspirations, and motivations. *CALL-EJ, 13*(1), 1–16. **http://callej.org/journal/13-1/Lockley_Promnitz-Hayashi_2012.pdf**

Milliner, B. (2015). Implementing a mobile-based extensive reading component: A report on student engagement and learning outcomes. In M. Iguchi, & L. Yoffe (Eds.), *JACET Summer Seminar Proceedings: Vol. 14. Mobile Learning In and Out of the Classroom: Balancing Blended Language Learner Training* (pp. 41–49).

Milliner, B. (2017). One year of extensive reading on smartphones: A report. *The JALT CALL Journal, 13*(1), 49–58. **http://doi.org/10.29140/jaltcall.v13n1.211**

Murray, A., & Blyth, A. (2011). A survey of Japanese university students' computer literacy levels. *The JALT CALL Journal, 7*(3), 307–318. **http://doi.org/10.29140/jaltcall.v7n3.124**

Nakaya, K., & Murota, M. (2013). Development and evaluation of an interactive English conversation learning system with a mobile device using topics based on the life of the learner. *Research and Practice in Technology Enhanced Learning*, *8*(1), 65–89.

Nishio, T., & Nakatsugawa, M. (2020). 'Successful' participation in intercultural exchange: Tensions in American-Japanese telecollaboration. *Language Learning &Technology, 24*(1), 154–168. **http://doi.org/10125/44714**

Oberg, A., & Daniels, P. (2013). Analysis of the effect a student-centred mobile learning instructional method has on language acquisition. *Computer Assisted Language Learning, 26*(2), 177–196. **http://doi.org/10.1080/09588221.2011.649484**

Ockert, D. (2013). Japanese JHS student ideal L2 selves: Gender differences that appear from a self-determination theory survey. *The School House, 21*(1), 12–22.

Ockert, D. (2014). The influence of technology in the classroom: An analysis of iPad and video intervention on JHS students' confidence, anxiety, and FL WTC. *The JALT CALL Journal, 10*(1), 49–68. **http://doi.org/10.29140/jaltcall.v10n1.167**

Ohashi, L. (2016). Taking English outside of the classroom through social networking: Reflections on a two-year project. In S. Papadima-Sophocleous, L. Bradley, & S. Thouësny (Eds.), *CALL communities and culture – Short papers from EUROCALL 2016* (pp. 345–350). **http://doi.org/10.14705/rpnet.2016.eurocall2016.586**

Paterson, R. (2014). iPad integration: Mobile learning issues in university classes. *Language Research Bulletin, 28*, 1–14.

Pegrum, M. (2015). Mobile learning, today and tomorrow. In M. Iguchi, & L. Yoffe (Eds.), *JACET summer seminar proceedings: Vol. 14. Mobile learning in and out of the classroom: Balancing blended language learner training* (pp. 4–11).

Sana, F., Weston, T., & Cepeda, N.J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education, 62*, 24–31. **http://doi.org/10.1016/j.compedu.2012.10.003**

Stockwell, G. (2007). Vocabulary on the move: Investigating an intelligent mobile phone-based vocabulary tutor. *Computer Assisted Language Learning*, *20*(4), 365–383. **http://doi.org/10.1080/09588220701745817**

Stockwell, G. (2008). Investigating learner preparedness for and usage patterns of mobile learning. *ReCALL, 20*(3), 253–270. **http://doi.org/10.1017/S0958344008000232**

Stockwell, G. (2010). Using mobile phones for vocabulary activities: Examining the effect of the platform. *Language Learning & Technology, 14*(2), 95–110. **http://llt.msu.edu/vol14num2/stockwell.pdf**

Stockwell, G., & Liu, Y.C. (2015). Engaging in mobile phone-based activities for learning vocabulary: An investigation in Japan and Taiwan. *CALICO, 32*(2), 299–322. **http://doi.org/10.1558/cj.v32i2.25000**

Thornton, P., & Houser, C. (2002). M-learning: Learning in transit. In P. Lewis (Ed.), *The changing face of CALL: A Japanese perspective* (pp. 229–243). Swets & Zeitlinger.

Thornton, P., & Houser, C. (2005). Using mobile phones in English education in Japan. *Journal of Computer Assisted Learning 21*. 217–228.

Toland, S., Mills, D., & Kohyama, M. (2016). Enhancing Japanese university students' English-language presentation skills with mobile-video recordings. *The JALT CALL Journal, 12*(3), 179–201. **http://doi.org/10.29140/jaltcall.v12n3.207**

Ushioda, E. (2013). Motivation matters in mobile language learning: A brief commentary. *Language Learning & Technology*, *17*(3), 1–5. **http://llt.msu.edu/issues/october2013/commentary.pdf**

Wang, J., Zou, B., Wang, D., & Xing, M. (2013). Students' perception of a wiki platform and the impact of wiki engagement on intercultural communication. *System, 41*, 245–256.

Wang, S., & Higgins, M. (2006). Limitations of mobile phone learning. *The JALT CALL Journal, 2*(1), 3–14. **http://doi.org/10.29140/jaltcall.v2n1.18**

Wang, S., Iwata, J., & Jarrell, D. (2016). Learning via mobile phones – Students' learning styles, needs, preferences, and concerns. *International Journal of Innovation and Learning, 19*(4), 431–443. **http://doi.org/10.1504/IJIL.2016.076797**

Wang, S., & Smith, S. (2013). Reading and grammar learning through mobile phones. *Language Learning & Technology*, *17*(3), 117–134. **http://llt.msu.edu/issues/october2013/wangsmith.pdf**

White, J., & Mills, D. (2012). Get smart!: Smartphones in the Japanese classroom. In A. Stewart & N. Sonda (Eds.), *JALT2011 Conference Proceedings*. JALT.

White, J., & Mills, D. (2014). Examining attitudes towards and usage of smartphone technology among Japanese university students studying EFL. *CALL-EJ, 15*(2), 1–15.

Wu, Q. (2015). Designing a smartphone app to teach English (L2) vocabulary. *Computers & Education, 85*, 170–179. **http://doi.org/10.1016/j.compedu.2015.02.013**

Wu, P-H. N., & Marek, M. W. (2016). Incorporating LINE smartphone affordances: Cross-cultural collaboration, willingness to communicate, and language learning. *International Journal of Computer Assisted Language Learning and Teaching, 6*(2), 56–73. **http://doi.org/10.4018/IJCALLT.2016040104**

Yamauchi, M., & Uchida, M. (2011, June 3–5). *Overcoming barriers to student engagement in using English online* [Conference Session]. JALTCALL2011, Kurume University, Kurume, Japan.

Yang, F., & Wang, S. (2014). Students' perceptions toward personal information and privacy disclosure in e-learning. *Turkish Online Journal of Educational Technology, 13*(1), 207–216.

Author's Bio

Dr. Edo Forsythe is the Head of the Department of English Language and Literature at Hirosaki Gakuin University. His research interests include MALL and specifically the use of smartphones in language learning.

https://doi.org/10.37546/JALTSIG.CALL2020.4

4

Enhancing Teacher Education Through a Lesson Study Application Built with WordPress

James M. Hall, Kodai Tamura, and Shuhei Kumagai, Iwate University

Abstract

A WordPress Lesson Study Application (LS APP) was developed for an English teacher education program at a Japanese national university for the following two purposes: (1) bridge the gap between teacher education classwork and school practice; and (2) enhance collaborative reflective practice between student-teachers. The first was accomplished through linking student-teachers' teaching practicums with theory and techniques learned in their university English teaching methodologies class using the tagging function of the LS APP. The second was accomplished through enabling students to give each other feedback on their lessons during their practicum using the LS APP. This paper describes how WordPress content management functions and two plugins were utilized to accomplish the two objectives. Then, it describes the lesson study process facilitated by the application. Lastly, the authors describe the extent to which the LS APP has linked theory and practice as well as the type of peer feedback student-teachers perceive themselves to have benefitted from. The goal of this paper is to describe how WordPress functionalities have been used to facilitate teacher development, and serve as a practical case study of using technology to enhance teacher education.

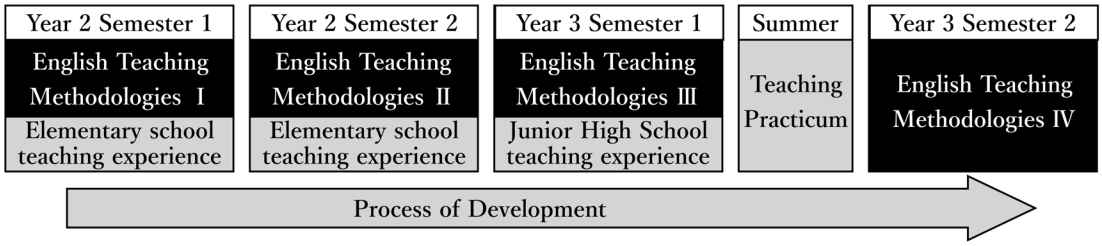
日本の国公立大学における英語教員養成課程において、WordPressを使用した授業研究アプリ（LS APP）は　(1)大学における講義と教育現場における実践の差をなくしていくこと、(2)実習生による共同の省察的実践を高めていくことの2つを目的に開発が進められてきた。1つ目の目的は、LS　APPのタグ付け機能を活用しながら、実習生による現場での授業実践と大学で学んだ理論や技術を結びつけることによって達成された。2つ目の目的については、授業実践が行われている最中にLS　APPを使用しながら参観している実習生がコメントし、互いにフィードバックしあうことによって達成された。本研究では、WordPressのコンテンツ管理機能及び2つのプラグインがどのような作用によって、これら2つの目的を達成しているかについて述べている。次にアプリを使用した授業研究の過程について、そして、実習生にとってどのようなピア・フィードバックが有益かだけでなく、LS　APPがどの程度まで理論と実践を結びつけているのかについて述べる。本研究における目標は、WordPress特有の機能性が教師としての成長を促進するためにどのように使用されているのか、そしてこの機能性が教員養成をよりよくしていくために、テクノロジーを使用した実践的事例研究としてどのように寄与するものかを述べることである。

Keywords: **Reflective practice, lesson study application, English teacher education, WordPress**

The first author works in the Faculty of Education at a national university in Japan and helps to oversee the English teacher education program. To get a first-level license (*isshu menkyo*) for teaching English in junior high school, student-teachers (STs) are required to take four English teaching methodology (ETM) classes, which will be called the ETM program hereafter. Figure 1 shows how the ETM Program mixes university class work with actual teaching experience. In ETM I to ETM III, STs have the opportunity to teach one lesson at an elementary or a junior high school to try out the theories and techniques learned in the ETM classes. For these lessons, they teach in groups of 2 to 4 people. Between ETM III and ETM IV, STs undergo their required teaching practicum for their license. In ETM IV, they reflect on their practicum and then complete their ETM classwork.

Figure 1

English Teaching Methodologies (ETM) Program



The ETM program takes a sociocultural approach to teacher development (Johnson, 2009). That is, it assumes that an ST's understanding of theory and practice will develop at the same time. According to Golombek and Johnson (2019), praxis is the dialectical unity of theory and practice in which "theory guides practice and practice shapes theory" (p. 25). In ETM, it is hoped that the praxis of STs will develop through classwork, practical experience, and STs attempting to make sense of their teaching experiences through collaborative reflective practice. Thus, the goal of ETM is to help STs become reflective practitioners who can develop their own theoretically sound and contextually appropriate teaching practices.

The first author designed a Lesson Study Application (LS APP) using WordPress to help STs develop into reflective practitioners. The LS APP enables STs to connect their teaching experiences with the learning areas in ETM classes and engage in collaborative reflective practice by giving feedback to one another on the lessons they teach. Not only do these purposes address the needs of the ETM program, but they also address two prevalent issues in English teacher education: making pre-service educational content relevant to actual teaching and encouraging collaboration in reflective practice. The literature review provides a brief background on the two issues and then discusses how WordPress can be used as an LS APP to address these issues. The authors investigated the extent to which the LS APP was able to help STs link their teaching experiences with ETM class content as well as the kind of feedback given on the LS APP that STs found beneficial. It shows tagging lesson study posts with ETM keywords has the potential to connect theory and practice, and how commenting on lesson study posts promotes collaborative reflective practice.

Literature Review

Issues in Teacher Education

It is generally accepted that there is a divide between the educational theory STs learn in the university classroom and the ways in which the STs teach at schools (Farrell, 2019). According to Baguley (2019), among the reasons for this divide are: theory that STs learned in class is not applicable to schools; STs have not learned strategies or techniques that would enable them to teach more flexibly; STs do not know how to teach to individual differences; STs are only familiar with linear lesson planning and struggle to make adjustments. Effective teaching depends on being able to make the most appropriate instructional judgements using one's expertise as well as situational understandings. Nutthall's (2007) longitudinal research on grade school students' classroom learning shows that there are no universal, one-size-fits-all methods for student-teachers to acquire: "The teaching that produces most learning in students varies from day to day, from class to class, and from time to time in the same class" (p. 24). In summary, to bridge the theory and practice divide, pre-service teacher education must help STs develop not only a repertoire of teaching techniques but a dynamic praxis which evolves as an ST's academic knowledge grows and pedagogical experiences increase.

The second issue is encouraging collaboration in reflective practice. In the field of English language teaching, reflective practice is considered a systematic way of analyzing and improving one's pedagogy (Farrell, 2015). Teachers plan a lesson, teach it, reflect on it, make improvements, and teach again. Reflection has become ubiquitous in the field of teacher education, but according to Mann and Walsh (2013), it is often carried out ineffectively. They argue that reflective practice is frequently done subjectively by individual teachers and much of their conclusions are reached without any kind of data-based support. Valid reflective-practice entails STs reaching their conclusions by being informed from data and the perspectives of others who observed their class.

Methods

The methods section first discusses how WordPress content management and plugins were utilized to create a Lesson Study Application that could address the two prevalent issues in teacher education: bridging the theory-gap divide and encouraging valid and collaborative reflective practice. Second, it explains the lesson study process. Third, it presents two research questions designed to investigate the extent to which the LS APP addresses these issues. Fourth, it describes the participants and main forms of data collected: lesson study tags and comments. Lastly, it discusses how the lesson study tags were categorized and the comments were coded.

Building a Lesson Study APP From WordPress to Support Teacher Development

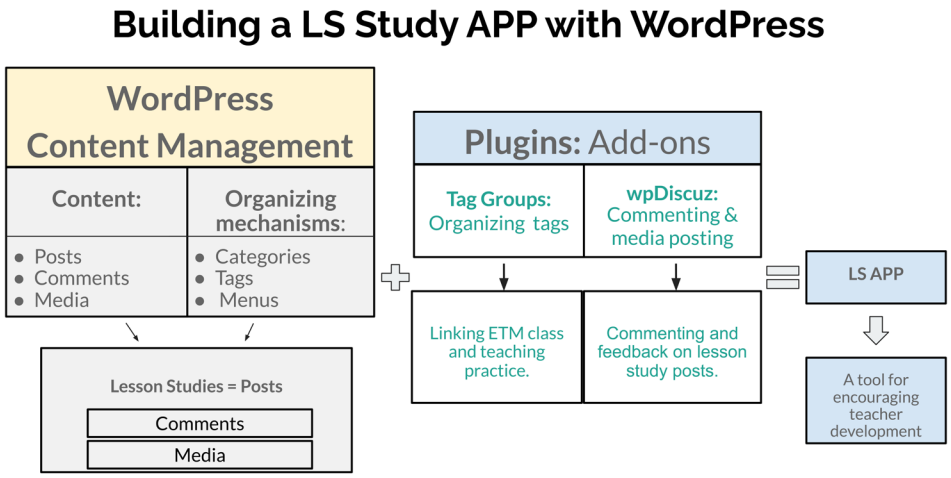
WordPress is a content management system (CMS) and popular website building platform powering about 32% of all sites on the internet (Messenlehner & Coleman, 2019). It is

well known as a blogging platform, and is used in education for ePortfolios, student blogs, classroom websites, and learning management sites (Edublogs, n.d.). WordPress can also serve as a platform for building web applications, which are defined as "software designed to perform a group of coordinated functions, tasks, or activities for the benefit of the user" ("Application Software," n.d.). In the case of the LS APP, the benefit of the coordinated functions is teacher development. These functions are explained more below.

The LS APP utilizes WordPress content management functions to organize the lesson study content. Plugins are used to link ETM learning areas with STs' teaching and to encourage collaborative reflection. The LS APP uses a total of 30 plugins for site security, site maintenance, styling, and user management. This paper focuses only on the plugins related to making ETM content relevant to classroom teaching experiences and encouraging collaborative reflection. The other plugins are discussed in detail in Hall (2019). Figure 2 provides an overview of WordPress content management for the LS APP and how two plugins perform the functions to facilitate teacher development on the LS APP.

Figure 2

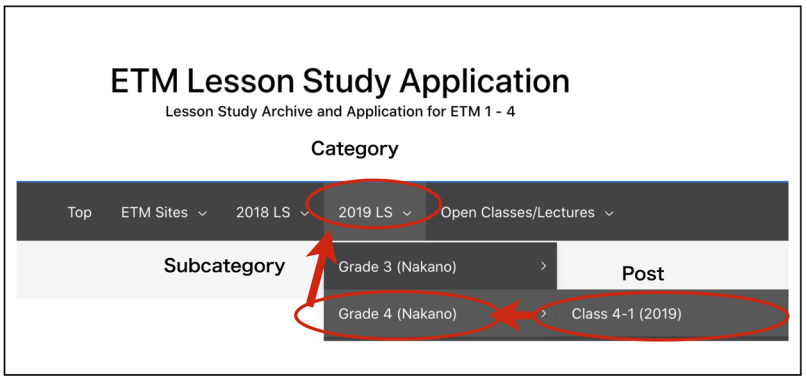
WordPress Content Management and Plugins Used on the LS APP



WordPress content management refers to the type of content appearing on the LS APP and the mechanisms for organizing it. The content of the LS APP is composed of posts, comments, or media. A post on the LS APP was a lesson study. The posts have media such as lesson plans, worksheets, video of the lesson, and pictures of the lesson as well as comments about the lesson attached to it. The content in the LS APP is organized by categories, menus, and tags. Categories are ways of grouping posts, and these groups of posts can be displayed in menus. The screenshot in Figure 3 shows the LS APP menu and the categories in which lesson studies were organized. For example, the lesson study post for class 4-1 was under the category *Grade 4,* which was a subcategory for *2019 LS*, or lesson studies conducted in 2019.

Figure 3

WordPress Menu Showing Lesson Study Categories



Post tags are further ways to organize content. A tag is a keyword which is attached to a lesson study and highlights its content. Figure 4 shows a lesson study post with the tags on the top. If users click on the tag, *Managing demonstrations*, they can see all lesson studies sharing the same tag as well as a description of the tag.

Figure 4

An Example of Tags as a Means of Organizing Content

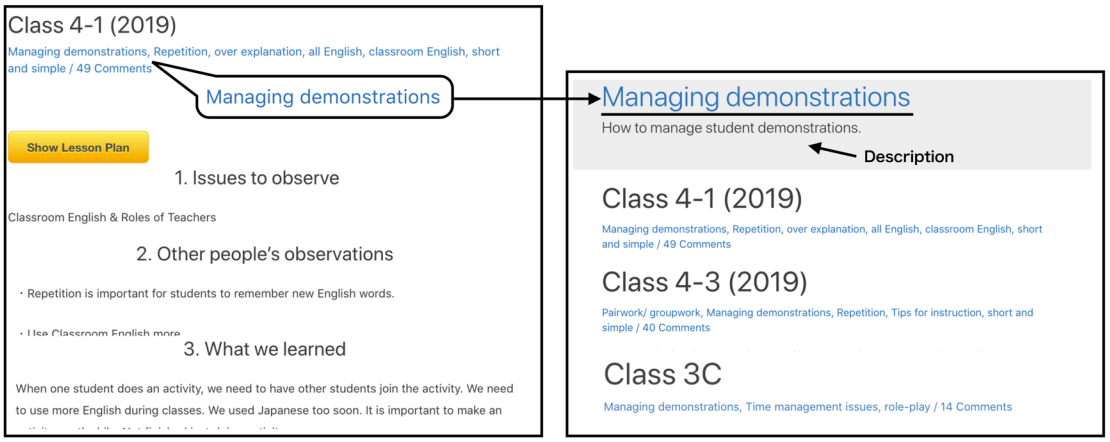


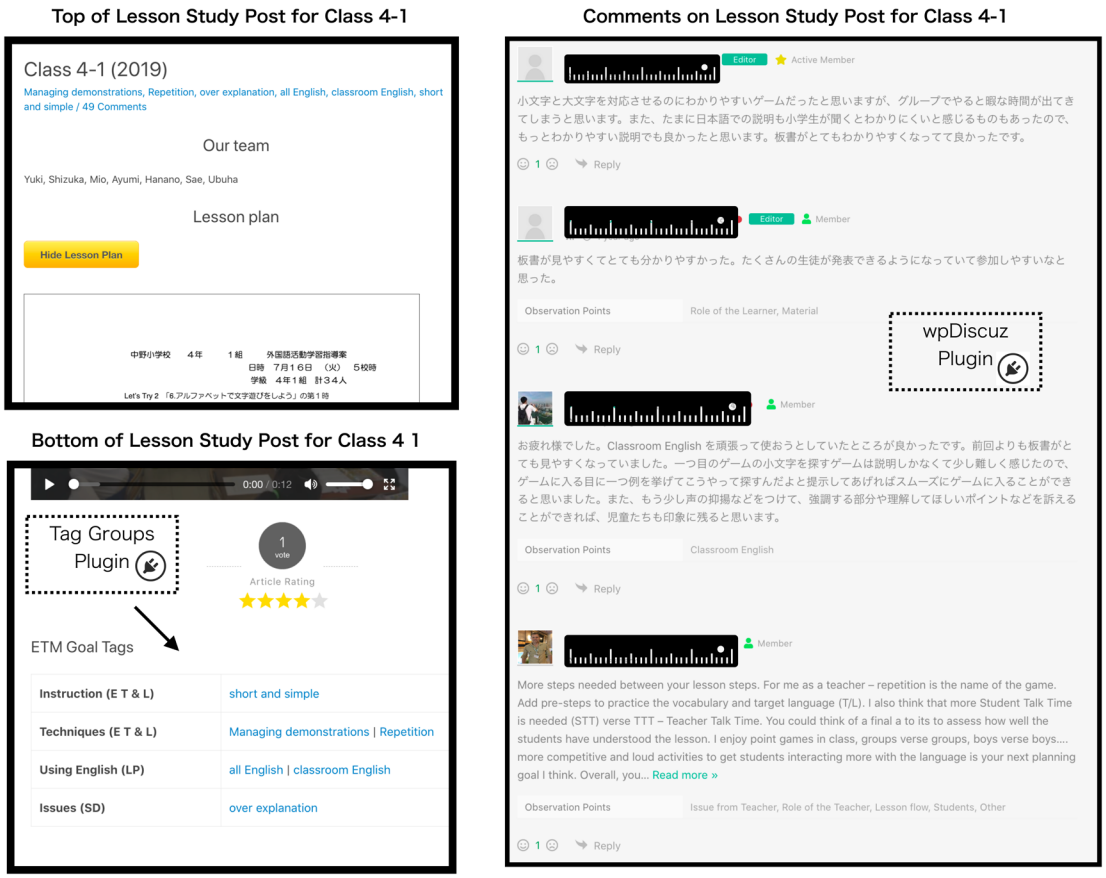
Figure 5 shows how the plugin, Tag Groups (Chatty Mango, n.d.), enables users to link ETM class content with their teaching experiences. Tags or keywords for the lesson study appear on both the top and bottom of the post. Tags shown on the bottom of the post link to ETM learning areas. For example, the tag *short and simple* is shown under the ETM secondary area of *instruction*.

Figure 5 also shows screenshots of comments on a lesson study post. The LS APP used the commenting plugin wpDiscuz (gVectors Team, n.d.) to facilitate valid and collaborative reflection. The screenshot shows users could choose an *Observation Point* for their comment. In addition, users could tag the practitioners or other observers in their comment encouraging them to read it and respond. Lastly, users could attach video or pictures to their

comments to substantiate their points. Thus, commenting enabled users to give and receive feedback on lessons as well as provide data to support their assertions. It was hoped that these functions would support valid reflective practice.

Figure 5

Tag Groups and wpDiscuz Plugin on a Lesson Study Post



Lesson Study Process

The lesson study process has five steps. Steps 1 to 3 involve planning, teaching, and receiving feedback. Steps 4 and 5 involve reflection. The first three steps from a lesson study together with screenshots from the application are shown in Figure 6. These steps show how feedback is facilitated on the LS APP making use of WordPress content management and the wpDiscuz plugin. In Step 1, a lesson study page or post is created. STs embed their lesson plan and write issues they want others to observe into the page. The *Issues to Observe* shown in the figure is classroom English. In the second step, a mock lesson is conducted and STs receive feedback from their classmates. Figure 6 provides an example of STs receiving feedback on their classroom English in the mock lesson. The feedback is written as a comment on the lesson study post. STs revise their class and then teach the lesson at a

school in Step 3. During the lesson, fellow STs, school teachers, and ETM instructors write their feedback into the LS APP. Step 3 shows a comment about STs' way of instruction and activity they conducted. In the comment form, users can choose a category for their feedback. They can also attach video or pictures to the comments. The comment shown in Step 3 has video attached.

Figure 6

Steps 1 to 3 of the Lesson Study Process

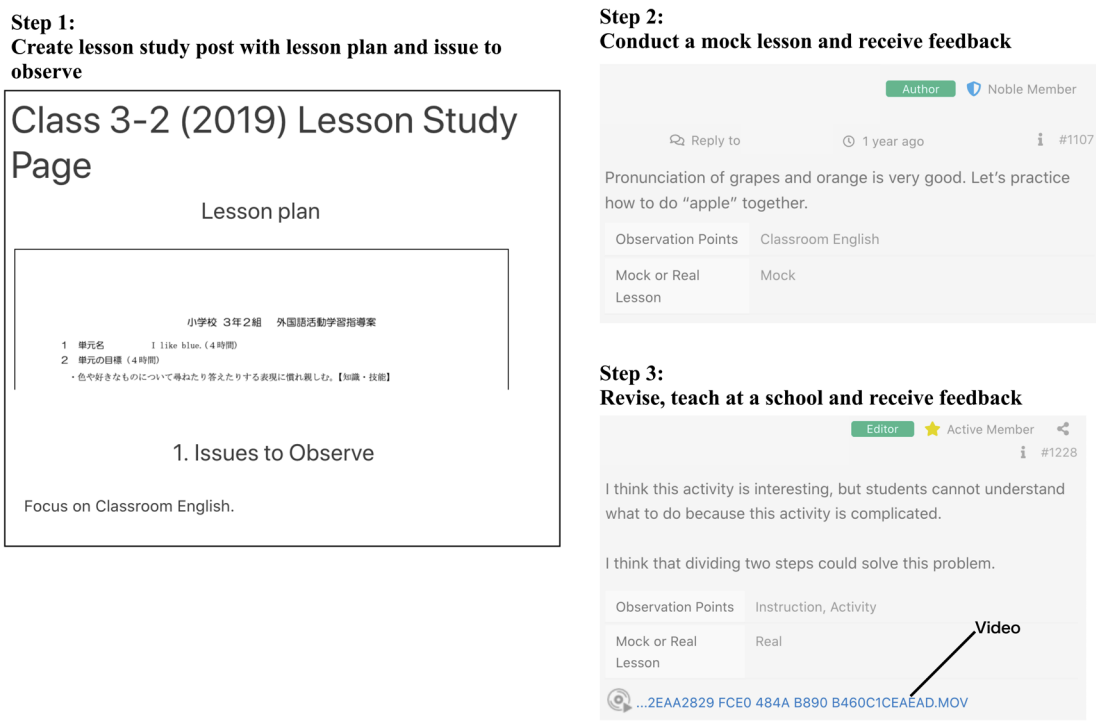
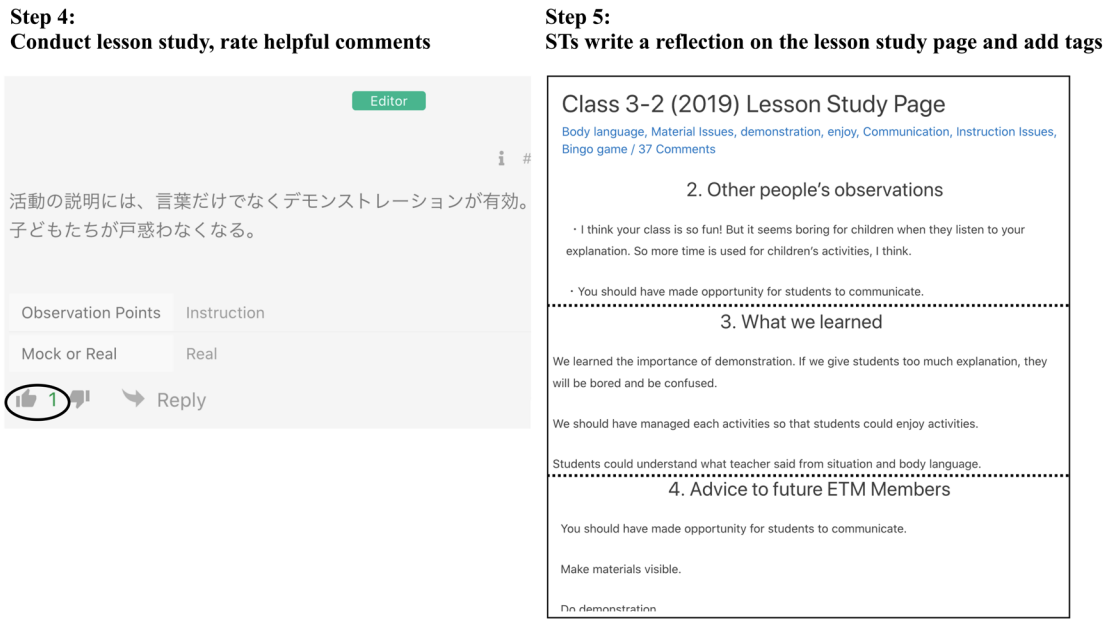


Figure 7 shows the steps related to reflection. In Step 4, a lesson study is conducted in which the comments are discussed. The screenshot for Step 4 shows a comment that was voted as beneficial by an ST. It was an observation about the elementary school students taking the lesson. STs vote comments as helpful by "liking" them. In Step 5, teachers write a reflection into their lesson study post and then write tags for the post. These tags are then related to ETM goals. Figure 7 shows that the STs summarized other people's observations as well as what they learned, and they gave advice to future members of the ETM program. They also wrote tags summarizing the content of their post. These tags were matched with ETM goals by the first author utilizing the Tag Groups plugin.

Figure 7

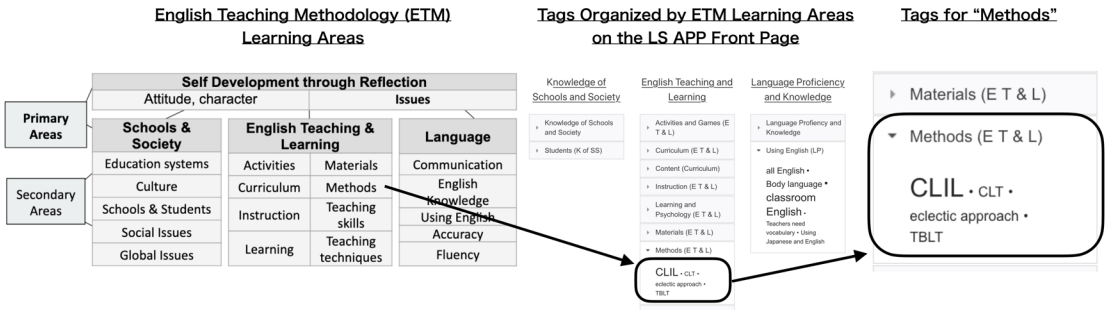
Steps 4 and 5 of the Lesson Study Process



Lastly, Figure 8 shows the learning areas of ETM I to ETM IV on the left. There are four primary areas: *Self Development through Reflection*, *Knowledge of Schools and Society*, *English Teaching and Learning*, and *Language Proficiency and Knowledge*. Each of the primary areas consists of secondary areas. The right side of the figure shows how tags are organized by ETM learning areas on the front page of the LS APP. For example, the figure shows that the secondary area *Methods* is a tag group. By clicking on *Methods*, users can see lesson study tags related to it.

Figure 8

ETM Learning Areas and Corresponding Tag Groups on the LS APP



Research Questions

Thus far, this paper has explained the goals of ETM, two issues in teacher education, and how an LS APP using WordPress was developed to address these two issues. The authors analyzed the lesson studies conducted in 2019 on the LS APP to answer the following questions:

1. To what extent does the LS APP bridge the divide between university classwork and teaching?
2. What kind of feedback on the LS APP did STs rate as being useful for their professional development?

Question 1 is clearly related to the first issue in teacher education, bridging the gap between theory and practice. For question 2, it was hoped that identifying the kind of feedback STs found useful would reveal the nature of ST collaboration as well as how STs perceived the benefits of engaging in collaborative reflective practice on the LS APP.

Participants and Data Collection

The data consist of 22 lesson study posts created in 2019, the tags written on the posts, and 1055 comments left on these posts. The comments were written by a total of 86 people, who were either STs currently taking ETM, graduate school students, or ETM instructors. Of the 1055 comments, 328 received a *like*, or were voted as beneficial by the participants. Participants were notified beforehand that their comments could be used for research purposes. When LS APP artifacts are shown, the names of the participants are hidden to protect their identity. It should be noted that participants wrote their comments with the understanding that their identity will be known to other members of their learning community, ETM program members.

Data Analysis

The tags written in the lesson-study posts were matched with ETM learning areas using the Tag Group plugin. It should be reiterated that the tags were written by the STs, but were matched to the ETM learning area by the authors. It was hoped that this would reveal what areas of learning in ETM were relevant to the STs' teaching. When lesson studies are conducted in the future, the authors plan for the STs themselves to match their tags with the ETM learning areas.

To consider the type of feedback the STs found useful for their professional development, the STs were asked to *like* (see Figure 7) any comments they found to be useful advice for improving their teaching. The 328 comments that the STs *liked* were analyzed. The 328 comments were coded inductively based on Kuckartz (2014). Firstly, the three researchers analyzed the comments or feedback given to STs during the classes and agreed on a set of codes. After that, the researchers coded the data individually, periodically comparing their codes on the feedback. Lastly, the researchers reviewed the comments together to ensure their agreement on the codes. A total of 15 codes for feedback were created. They will be

explained below. The authors gave the data one further classification: the purpose of the feedback. That is, was the feedback intended to give praise, critique the ST, or provide an observation? In summary, the authors analyzed the feedback the STs considered to be useful. The initial code revealed the content of the feedback and the second code revealed its purpose.

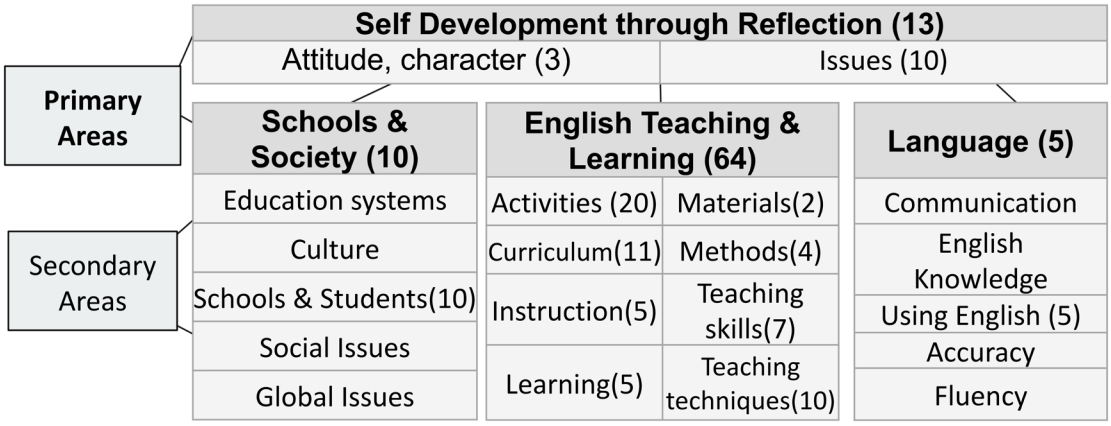
Findings

Bridging ETM Class Content and ST Teaching on the LS APP

Figure 9 shows the number of lesson study tags related to the ETM content areas. If a secondary area had associated tags, the names of some sample tags were written into the figure. The most frequent tags on secondary areas are those related to *Activities*. A goal of ETM is to learn various activities that STs can use in their teaching practice. An example of activities STs learned about in ETM and then actually used were role play and bingo. The second most frequent secondary area was *Curriculum*, this is largely related to lesson planning, creating learning objectives, assessing these learning objectives, and considering the place of communication in the curriculum. *Techniques* and *Issues* were the third most frequent secondary area. *Techniques* refer to ways to manage students or conduct practice activities such as chants. *Issues*, as mentioned earlier, refers to problems the STs encountered.

Figure 9

Number of Lesson Study Tags for ETM Learning Areas



On the other hand, secondary areas related to academic concepts such as *Learning* and *Methods* had lower frequencies. It should be noted that when the STs do a lesson, they are preoccupied with carrying out their lesson plan, conducting activities, and encouraging student participation. Academic concepts covered in ETM such as *Noticing* as well as *Task Based Language Teaching* (TBLT) are not of immediate relevance to STs. This reflects research findings that have found novice teachers to be more focused on the here and now or carrying out individual lessons (Richards et al., 1998).

Overall, the matching of tags and learning areas shows that STs were able to apply a

variety of practical knowledge they learned in their ETM classes to their once-a-semester teaching practice. However, academic concepts learned in ETM do not seem to be as relevant to the teaching practicum.

Types of Feedback Given on the LS APP that STs Found Useful

Table 1 shows a description of the feedback codes. In general, most of the feedback STs found helpful focused on their actions during the class. This kind of feedback was related to how STs tagged lesson study posts, which covered practical areas of the ETM curriculum.

The two most common types of feedback were about the activities done in the class and the STs' way of instruction. *Activities* expressed the observers' impressions, and often discussed how the STs could modify the activity to make it easier for students to participate in. An example of such feedback is as follows: "I think that the message game is a little difficult for children, especially, the two sentences. They are difficult even for us. So, I think either the activities or sentences have to be changed" (Edited for clarity by the authors).

Table 1

Codes and Frequencies (N=328)

Feedback Code	Description	Frequency
Activities	About a class activity.	65
Instruction	Critiquing the STs' instruction.	58
Overall	About the entirety of the lesson.	44
Using English	About ST's use of English.	35
Students	An observation focusing on the students.	28
Picture	A picture is attached to a comment with a caption.	22
Lesson flow	About the overall lesson flow.	19
Materials	About any materials used for teaching.	13
Video	A video is attached to a comment with a caption.	10
Attitude	About the attitude exhibited by the STs.	7
Objectives	About the goal or objective.	4
Philosophy	About the ST's teaching philosophy.	3
Self-reflection	Self-reflection of the ST.	2
Lesson planning	About lesson planning.	2
Encouragement	Intended to encourage the ST.	1
N/A	Cannot be given a code.	15

Instruction was feedback on how an ST demonstrated or explained language or a classroom procedure. An example is as follows: "When you explain how to play the Leap Frog Game, you should tell students the procedure of the game (e.g. "First, read the cards. Second, etc...)." Often, *Activity* and *Instruction* feedback would co-occur: the commenter would discuss both an activity as well as how the ST demonstrated or explained the activity.

The following comment has both elements: "Activities were very interesting! You all have to use more gestures to explain because some of the words were difficult for students" (Edited for clarity by the authors).

Although the above comment was coded as *Activity*, it could have also been coded as *Instruction*. Therefore, rather than concluding that the most frequent comment was about *Activity*, it is most appropriate to say that feedback on activities and/or instruction was most frequent.

Using English referred to the STs' ratio of L1 to L2, possible mistakes the ST made, suggestions about language they can use, or complimenting their English usage. The following comment is an example: "It was good. Classroom English was good too, but a little difficult for elementary school students, I thought. You can use easier English and speak more slowly with many gestures and feelings" (Edited for clarity by the authors).

Feedback coded as *Activities*, *Instruction*, *Using English*, *Materials*, *Attitude*, *Philosophy*, *Lesson-planning*, and *Encouragement* focused exclusively on the ST's behavior, actions, or work. Other comments, however, served to inform the ST about other aspects of the class. Comments coded as *Students* focused on the behaviors and characteristics of students. Their purpose was to share observations about the students that the STs might not have noticed or recommend better ways to manage the students. The following comment was coded as *Students*: "Almost all the students understood the rules and were diving into the activity. It looked like a lot of fun" (Translated by the authors).

The *Picture* and *Video* codes show that feedback was not only in the forms of written comments. Observers posted audio-visual material as comments to show how STs were teaching, students' actions, classroom layout, or student work. Thus, often, the multimedia comments served to inform STs about what was happening in the classroom.

Overall was the third most-frequently voted of all the feedback codes. While comments like *Activities* tended to be written during a lesson, *Overall* tended to be a cumulative set of observations written at the end of the lesson or after it was finished. Many of the overall comments were organized, starting by praising and, after that, critiquing an ST's class. They also were written from the perspective of both the teachers and students. The one individual comment that received the most votes from STs as being helpful (5 votes) was classified as *Overall* it reads as below:

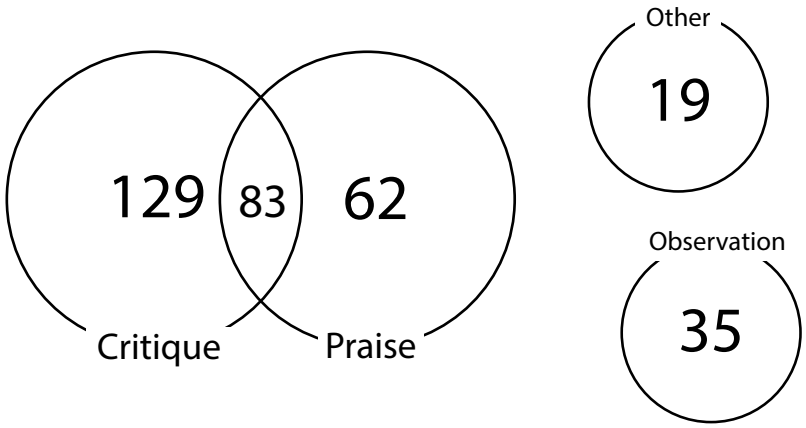
> *Good job on your lesson! I liked how you taught the class with a smile. In terms of lesson flow, I thought it was a little sudden how you went from presenting the goal to practice. Also, the goal was not so clear. The practice activities were connected with the main activity in that students practiced many words they would need to use. For a good lesson flow, you need to make it clearer what you expect the students to get out of the activities and encourage students to think "I want to get better at this; I want to learn how to do this!" during the activities* (Translated from Japanese by the authors)

Figure 10 is a Venn diagram which shows how the purpose of the different feedback comments were coded. This diagram shows that *Critique* comments had the highest frequency of helpful votes, or 212, which constituted two-thirds of all feedback voted as helpful. It is noteworthy that the STs found negative feedback to be helpful. However, it should also be noted that negative feedback was often combined with praise. There were 83 comments

that were coded as both *Critique* and *Praise*. The implications are that STs are receptive to negative feedback, but further inquiry needs to investigate why and the characteristics of negative feedback found helpful by STs.

Figure 10

A Venn Diagram of the Purposes of the Feedback (N=328)



Discussion

LS APP as a Means of Bridging the Theory and Gap Divide

This section discusses the potential of the LS APP to bridge the gap between teacher education and the actual classroom as well as some issues. The findings show that the LS APP could be used as a tool to gauge which ETM areas were relevant to the STs' teaching practice. This can provide opportunities to improve the course program. By knowing which ETM areas are being overlooked in the lesson studies, the instructors can reconsider how they are introduced in ETM or ask STs to consider these areas in Steps 4 and 5 of the lesson study process.

As mentioned earlier, although the STs were writing the tags into their lesson study posts, it was the first author who matched the tags with ETM learning areas. Because one of the purposes of the LS APP was to help students develop their own praxis, or theory-informed-practice, this was a major issue. In future uses of the LS APP, the STs themselves will match their tags with the learning areas. Farr and Riordan (2017) argue that novice teachers need support in being able to relate theory and practice, it does not come naturally. Therefore, the tagging feature will not serve its purpose without guidance. One possible way for supporting STs is having them tag their lesson study posts in groups and discuss the rationale for the tags that they chose. Another way is to have STs read previous lesson study posts and discuss the relationships between the tags on the post and the content.

Another issue was not in the tags themselves but rather in the lesson study post content. Some of the reflections in the Lesson Study pages need to be written in more detail so that the reader can better understand the concepts represented by the tags. Sometimes, a tag

was barely discussed in its lesson study post. Overall, tagging shows the potential of the LS APP tagging system to help STs reflect on their praxis, but use of the LS APP should be improved to actualize this objective.

LS APP as a Means of Encouraging Collaborative Reflective Practice

In this section the authors will discuss both the potentials and limitations of the LS APP for encouraging valid collaborative reflective practice. The potential of the LS APP lies in the volume of feedback that was communicated as well as the type of data that was uploaded. It was discussed at the beginning of the paper that one of the issues with reflective practice is that reflection tends to be subjective and solitary. The STs were able to learn the perspectives of a number of different observers from reading the comments. Furthermore, tangible data such as pictures and video could serve as evidence to support the observers' assertions. Thus, overall, it can be concluded that the STs were able to see their lesson through the eyes of others.

Another potential is the systematic nature of the process. All ETM members follow the five steps of the LS Process, in which they receive feedback at two different steps and reflect at the end of the process in Step 5. They also have at least three opportunities to undergo this process in the ETM Program. The LS APP puts this process online as it archives the lesson study posts, comments, and media given in each stage. Therefore, the LS APP supports ETM members long-term systematic reflection over the span of 2 years.

The last potential is for making university teacher education more practical. Collecting lesson study comments over time can reveal the kinds of practical issues that STs are likely to have when teaching. These practical issues can, in turn, be incorporated into the teacher education curriculum.

The primary limitation is the challenge in giving feedback during a lesson. Previous research on fieldwork in education has shown that taking notes while observing a class is an acquired skill (Hall, 2017). Furthermore, writing observations into the LS APP interface using tablet or phone, a medium for classroom observation to which STs might not be accustomed, is even more challenging than conventional notetaking (Hall, 2019). It is the authors' experience that for some STs, and even some instructors, giving feedback using a device during a lesson is not possible. One way to address this issue is to give STs more opportunities to practice writing comments before the lesson studies. Another way is to explicitly give STs the option of writing comments at the end of a lesson if they prefer to concentrate solely on observation during the lesson.

Conclusion

The primary purpose of this paper has been to describe the development of the LS APP. The secondary purpose has been to show the type of findings that can be gathered through collecting data from the application. This paper first introduced two issues in foreign language teacher education: 1) bridging the gap between what is taught in university language teacher education programs and actual school classrooms, and 2) promoting valid and collaborative reflective practice. It then showed how WordPress content management

functions and plugins can be used to create a Lesson Study APP to address these issues. First, it showed how post tags with the Tag Groups plugin can be used as a means to bridge the gap between the university and school classroom and improve teacher education curricula. It also showed how utilization of the wpDiscuz plugin enabled users to provide feedback addressing specific issues and provide support for their observations by attaching videos or pictures.

When analyzing the tags left on lesson study posts, it was discovered that ETM members tended to link their teaching experiences more with developing skills in instruction and learning how to do activities and less with concepts learned in the ETM class such as noticing. This is to be expected given that ETM members are experiencing teaching for the first time and their immediate needs are to develop the practical skills necessary to do a lesson. Linking classroom instruction with theory does not come naturally, and STs will need more guidance on this. It can be said that the LS APP can be used as a tool to help STs link classroom instruction with concepts, but that they will need help in learning how to use the tool.

Through analyzing the votes on comments, the authors were able to ascertain that ETM members found feedback on how they conducted activities, their instruction, and English to be beneficial. 212 of the 328 comments found helpful contained some form of negative feedback. There were two ways in which ETM members wrote feedback to one another: commenting while observing as well as writing a long comment after the lesson. The latter, which was categorized as *overall*, had a high frequency of likes. Writing one's comments while observing a class is an acquired skill which requires some training.

This study has several limitations. First, the authors analyzed the type of feedback STs found beneficial through their votes, but a questionnaire asking STs about their experience giving and receiving comments could have confirmed the STs' preferences. Second, a primary purpose of the LS APP is to encourage STs to link ETM class concepts with their experiences. However, it was the authors who actually did this. At the end of 2020, the authors plan to have STs link their experiences with ETM learning areas in the LS APP and answer a questionnaire about using the tag groups plugin. Lastly, analysis of the comments show that the authors need to better promote feedback and reflection that goes beyond activities and instructional techniques. The LS APP is a tool with potential, but the quality of the teacher education program and the guidance provided to STs will determine its success.

References

Application Software. (n.d.) In *Wikipedia*. Retrieved September 5, 2020 from **https://en.wikipedia.org/w/index.php?title=Application_software&oldid=974356953**

Baguley, N. (2019). 'Mind the gap': Supporting newly qualified teachers on their journey from pre-service training to full-time employment. In S. Walsh & S. Mann (Eds.), *The Routledge handbook of English language teacher education* (pp. 125–137). Routledge. **https://doi.org/10.4324/9781315659824-10**

Chatty Mango. (n.d.). Tag Groups – WordPress Plugin. *Chatty Mango*. Retrieved September 4, 2020, from **https://chattymango.com/tag-groups-premium/**

Edublogs. (n.d.). *Edublogs – free blogs for education – Blogs and websites for teachers, students, and schools*. Retrieved September 4, 2020, from **https://edublogs.org/**

Farr, F., & Riordan, E. (2017). Prospective and practising teachers look backwards at the theory-practice divide through blogs and e-portfolios. In T.S.C. Farrell & T. Stewart (Eds.), *Preservice teacher education* (pp. 13–26). TESOL Press.

Farrell, T.S.C. (2015). *Promoting teacher reflection in second language education*. Routledge.

Farrell, T.S.C. (2019). Reflective practice in L2 teacher education. In S. Walsh & S. Mann (Eds.), *The Routledge handbook of English language teacher education* (pp. 25–37). Routledge. **https://doi.org/10.4324/9781315659824-4**

Golombek, P.R., & Johnson, K.E. (2019). Materialising a Vygotskian-inspired language teacher education pedagogy. In S. Walsh & S. Mann (Eds.), *The Routledge handbook of English language teacher education* (pp. 25–37). Routledge. **https://doi.org/10.4324/9781315659824-3**

gVectors Team. (n.d.). *wpDiscuz – The best WordPress comment plugin. WpDiscuz – WordPress comment plugin*. Retrieved September 6, 2020, from **https://wpdiscuz.com/**

Hall, J.M. (2017). Methodological lessons learned from conducting reflexive linguistic ethnography. *Doing research in applied linguistics*, *2*, 77–93.

Hall, J.M. (2019). Negative and positive aspects of using WordPress to facilitate reflective practice in pre-service teacher education. In R. Chartrand & E. Forsythe (Eds.), *AI and machine learning in language education* (pp. 18–32). JALT CALL SIG. **https://jaltcall.org/wp-content/uploads/2019/06/AI-and-Machine-Learning-in-Language-Education.pdf**

Johnson, K.E. (2009). *Second language teacher education: A sociocultural perspective*. Routledge.

Kuckartz, U. (2014). *Qualitative text analysis: A guide to methods, practice and using software*. SAGE.

Mann, S., & Walsh, S. (2013). RP or "RIP": A critical perspective on reflective practice. *Applied Linguistics Review*, *4*(2), 291–315. **https://doi.org/10.1515/applirev-2013-0013**

Messenlehner, B., & Coleman, J. (2019). *Building web apps with WordPress: WordPress as an application framework*. O'Reily.

Nutthall, G. (2007). *The hidden lives of learners*. NCZER Press.

Richards, J.C., Li, B., & Tang, A. (1998). Exploring pedagogical reasoning skills. In J.C. Richards (Ed.), *Beyond training* (pp. 86–102). Cambridge University Press.

Authors' Bios

James M. Hall is an Associate Professor at Iwate University and is involved in pre-service English teacher education. He has taught in Japan for almost 20 years. His research

interests include ICT in English teacher education and ethnography for understanding teacher development.

Kodai Tamura is currently a graduate student at Iwate University Graduate School of Education. He began to use a Windows 95 laptop when he was four years old. He currently majors in English education, and his research interests are CALL, lesson study, and the possibility of using ICT in public schools of Japan.

Shuhei Kumagai is currently a graduate student at Iwate University Graduate School of Education. He has specialized in English education for 6 years and has experience teaching at a Japanese high school and a Nepali school. He is also interested in deaf education to empower deaf and hard of hearing people.

https://doi.org/10.37546/JALTSIG.CALL2020.5

5

Exploring EFL Student Use of Digital Backchannels During Collaborative Learning Activities

Peter Ilic, University of Aizu

Abstract

This chapter highlights several findings related to the learners' use of digital communication channels during online collaborative activities. The term *digital backchannel* is used to imply that there are two channels of communication operating simultaneously during collaborative activities. The predominant digital channel is that of the online content management system controlled by the instructor and accessed in the target language, English. The secondary channel of digital communication (backchannel) is that of the external personal social network systems (SNS) that students employed to interact with group members and others. The researcher collected qualitative and quantitative data on learner interactions within a yearlong series of language learning activities through internet logs and interviews. The students employed digital backchannels with increasing frequency throughout the study period, even though there was an initial reluctance to use a communication channel, they considered very private for public educational activities. These digital backchannels were primarily mobile-based SNS. The students reported that the use of these backchannels increased in use over the study period and led to increased peer communication and networking. Also, students' language use changed when moving between these primary and secondary communication channels, with L1 being used for the main-channel and L2 used for the backchannel.

この論文は、オンラインの協働作業中の学習者によるデジタルコミュニケーションチャネルの使用に関連するいくつかの調査結果に焦点を当てています。「デジタルバックチャネル」という用語は、協働作業中に2つの通信チャネルが同時に動作することを意味します。主なデジタルチャネルは、インストラクターが制御し、ターゲット言語である英語でアクセスできるオンラインコンテンツ管理システムのチャネルです。デジタルコミュニケーションの2番目のチャネル（バックチャネル）は、学生がグループメンバーや他の人々とのやり取りに使用した外部のパーソナルソーシャルネットワークシステム（SNS）のチャネルです。私は、インターネットログとインタビューを通じて、1年にわたる一連の言語学習活動内の学習者の相互作用に関する定性的および定量的なデータを収集しました。学生は、公的教育活動のために非常にプライベートであると考えられていた通信チャネルを使用することには当初抵抗があったにもかかわらず、調査期間を通じて増加する頻度でデジタルバックチャネルを利用しました。これらのデジタルバックチャネルは、主にモバイルベースのSNSでした。学生は、これらのバックチャネルの使用は、学習期間中に使用が増加し、ピアコミュニケーションとネットワーキングの増加につながったと報告しました。また、これら

のプライマリコミュニケーションチャネルとセカンダリコミュニケーションチャネルの間を移動すると、生徒の言語使用が変更され、L1はメインチャネルに使用され、L2はバックチャネルに使用されます。

Keywords: **digital backchannels, EFL, collaborative learning, mobile learning**

This paper highlights several findings related to the learners' use of digital communication channels during online collaborative activities. The term *digital backchannel* (Yngve, 1970) is used to imply that there are two channels of communication operating simultaneously during collaborative activities. The predominant digital channel is the instructor's online content management system and accessed in the target language, English. The secondary channel of digital communication (backchannel) is that of the external personal Social Network System (SNS) that students employed to interact in any language they choose because the instructor did not monitor these channels.

The researcher collected qualitative and quantitative data on learner interactions within a yearlong series of language learning activities through internet logs and interviews. The activities were online discussions and accessible through a range of mobile and non-mobile devices to allow the method participants found most agreeable. The participants were studying English at a four-year private university in Tokyo, Japan. Commonly used digital backchannels were SNS, such as Twitter and Line, which were drawn into the collaboration. The initial reluctance to use these channels of communication for homework was overcome as the students recognized their value as a digital backchannel for private information exchange independent from the main channel, a Moodle site. The digital backchannels allowed communication in their L1 of Japanese, which appears to have reduced their motivational barriers to the homework by providing support for their public use of English by reducing the potential for embarrassing mistakes, among other reported reasons. The paper includes a review of relevant literature, a summary of the results, a discussion of the findings, and a conclusion with limitations and possibilities for furthering the study.

Literature Review

The term *backchannel* was originally designed to imply two communication channels operating simultaneously during a conversation. The predominant (front) channel is that of the primary communication flow. In linguistics, the secondary (back) channel of communication is a verbal and non-verbal listener response serving a meta-conversational purpose (Yngve, 1970) that may vary in frequency across cultures (White, 1989). This second channel improves the communication process by augmenting the primary channel of information with various mannerisms, actions, gestures, and verbal expressions (Harry et al., 2009). The linguistic Backchannel Output Hypothesis suggests that backchannels may facilitate the fluency of beginner English learners' fluency during oral tasks (Wolf, 2008).

Although historically, the phrase *backchannel* has referred to these verbal utterances and non-verbal body language, its use has expanded. First, the term *digital backchannel* came to refer to synchronous non-verbal, real-time communication, which does not interrupt a presenter or event (McCarthy & Boyd, 2005). Later, it was used to identify asynchronous

microblogging involving the posting of digital content, such as text, pictures, links, short videos, or other media to web-based sharing services (McNely, 2009).

Researchers have investigated synchronous and asynchronous digital backchannels and their role in group interactions, such as one-on-one classroom chats (Cogdill et al., 2001) or public chat backchannels in physically-shared spaces such as an academic conference (McCarthy & Boyd, 2005).

Asynchronous backchannels are not as constrained by the limits of time or space, so are considered a non-disruptive collaborative activity that increases participation and interactions among students (Toledo & Peters, 2010) by allowing all students an equal opportunity to respond to a topic and more time to think and edit, so increasing the chance of thoughtful responses (Birch & Volkov, 2007; Branon & Essex, 2001; Hanson-Smith, 1997; Kitade, 2008; Ortega, 1997). One reason for these improved responses may be that they have more time to process input (Abrams, 2003) and reflect on what they want to express (Althaus, 1997). By leveraging the social communication, connectivity, and heightened multi-tasking skills associated with the everyday lives of students, it is possible to free up the valuable face-to-face interaction time (Jarrett & Devine, 2010; Williams, 2000) for more comprehensive in-class discussions that may positively influence subject matter reflection (Donnelly, 2016). For example, Twitter microblogging as a backchannel technology affords mobility through SMS messaging capabilities (Honeycutt & Herring, 2009), allowing users to engage in collaborative research anytime and anywhere. These microblogging systems also provide near-instantaneous responses with which students have become accustomed (Toledo & Peters, 2010), and information transfer through different media modes (Schick, 2005) such as text and video.

Digital backchannels provide a persistent online space that supports discussions (Carpenter, 2015), so they are ideal for developing communities of practice within higher education (McNely, 2009; Yardi, 2006). Backchannel communication provides support through peer-networking (Hennessy et al., 2016), and it improves the ability to learn from peers in a positive manner (Rogoff et al., 2004) by providing a digital space where students have the freedom to direct discussion that is relevant for their learning purpose in order to create their knowledge (Yardi, 2006). Participants in backchannels tend to be more sociable (McNely, 2009) and make more new connections (Brooke, 1987; Toledo & Peters, 2010). As Erickson and Kellogg (Erickson & Kellogg, 2000) have argued, digital backchannels are mechanisms for creating social proxies that provide a resource for group interactions online that, in turn, helps to build common ground through co-presence and visibility (Clark & Brennan, 1991).

Digital backchannels provide students a greater sense of ownership over learning that shifts the control over learning to students resulting in increased engagement (Camiel et al., 2014). They engage marginalized shy and introverted students by providing a more comfortable mode of communication that is a less threatening way to present ideas (Camiel et al., 2014; Carpenter, 2015; Krishnan & Poleon, 2013; Toledo & Peters, 2010). A digital backchannel chat platform has been shown to promote students' engagement in large English as a Foreign Language (EFL) classes by transferring the side conversation to the forefront (Harunasari & Halim, 2019).

Students who used online discussion backchannel communication have been shown

to have higher critical thinking levels than students who do not use online discussion (Rathakrishnan et al., 2017). Discourse outside front-channel instruction often involves questioning ideas, exposure to alternative points of view (Chen & Looi, 2007), building ideas that may be in conflict with accepted ways of thinking and acting (Brooke, 1987), and introducing unofficial references without disrupting the main-channel interaction (Cogdill et al., 2001), since one person cannot easily dominate the discussion (Branon & Essex, 2001; Ortega, 1997).

A digital backchannel archive can facilitate student assessments or reviews (Carpenter, 2015) and enable teacher self-assessment (Yardi, 2006) and monitor group dynamics to improve future classes (Krishnan & Poleon, 2013; Yardi, 2008; Yates et al., 2015).

Digital backchannels have been utilized to connect synchronously in large classes and venues to support various forms of student-to-lecturer and student-to-student interaction to make interactions in large classes similar to those occurring in small classes (Beatty et al., 2006; Donnelly, 2016). These synchronous interactions do not form a single conversation but instead multiple monologues with a few dialogues between users (Ross et al., 2011). In the classroom, synchronous digital backchannels can give students greater public anonymity when names are not attached to in-class posts while still being privately accountable to the instructor through backchannel logs (Carpenter, 2015). So, students who feel less competent may be encouraged to express their opinions aloud in these backchannel discussions (Bry & Pohl, 2017; McNely, 2009). They also emphasize active listening and informal learning (Toledo & Peters, 2010). One well studied synchronous technology is the audience response systems that have been used to support lecturers' question asking (Caldwell, 2007; Fies & Marshall, 2006; Kay & LeSage, 2009; Lantz, 2010).

Researchers have noted that there may be issues with distraction, which can be supported by the Cognitive Load Theory (Sweller, 1994). Students have always been subject to distractions during class, but in today's connected world, the possibilities for distractions increase, with some suggesting the term *continuous partial attention* describes student cognitive ability to pay attention (Yardi, 2006). Students may process content information on a more superficial level while they divide their attention across multiple domains simultaneously (Hembrooke & Gay, 2003). Researchers have also highlighted other negative issues related to digital backchannels, including rude content, ingroup versus outgroup conflicts, and effects on main-channel participation (Cogdill et al., 2001; McCarthy et al., 2004; McCarthy & Boyd, 2005; Yankelovich et al., 2005). Yardi (2006) recommends that the implementation of backchannels must vary across different contexts and domains and that a backchannel etiquette needs to be developed (Yardi, 2006).

Methodology

The study design was a case study adopted for one academic year to gain a deeper understanding of the processes and outcomes of completing collaborative learning activities through mobile devices by four Japanese university undergraduate EFL classes on translation. One case study group was formed from each class to make four case study groups with between five to eight members. All activities and environmental factors remained constant across all the case study groups, and participation in the study was voluntary. Group 1

contained five girls and two boys, group 2 contained eight girls, group 3 contained six girls, and group 4 contained six girls. There was no restriction on the type of device allowed to access the course Moodle site. The data collection, content, and procedures for each group were identical, and all the interviews took place in the same location with a single interviewer and were of approximately equal length.

The collaborative activities used in this study were group collaborative learning activities within weekly online modules. These activities consisted of homework questions related to translation issues encountered when moving between the Japanese and English languages. These made the students aware of specific known difficulties in language translation between these languages to encourage authentic use of their L2. This would serve the dual goal of content knowledge development and L2 practice. These activities required collaboration to complete so they would stimulate some discussion in the L2.

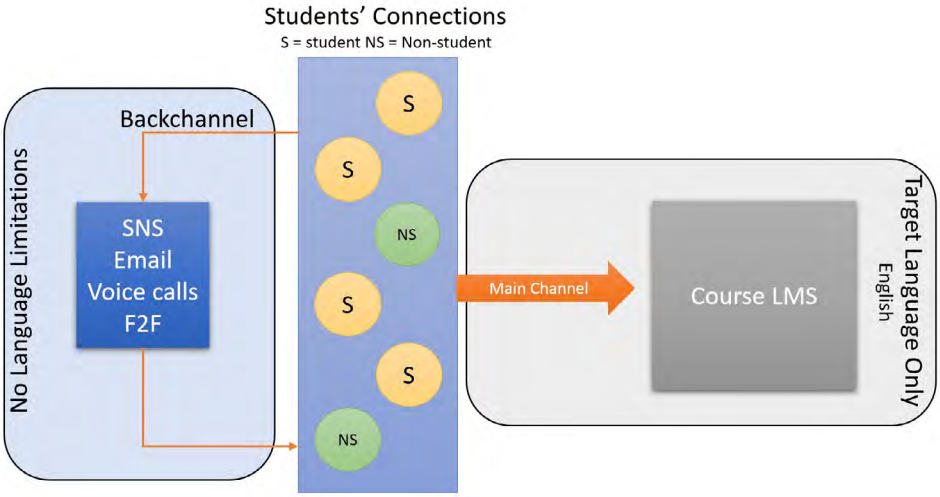
Student online website log data, weekly e-journal reports, and pre- and post-study face-to-face interviews comprised the principal sources of data. The first stage of data analysis was content analysis, where data was coded then categorized into themes. An inductive form of thematic analysis coding (Ezzy, 2002) was adopted to identify themes or concepts in the data, build a systematic account of what has been observed, identify emergent theory, and highlight issues and problems not anticipated. A single researcher performed this coding.

Results

This chapter highlights several preliminary findings related to the learners' use of digital communication backchannels channels during online collaborative activities. Figure 1 is an outline of the digital communication channels that were mapped during the case study research. Here the main channel of communication is represented on the right by the Moodle site. Here the language of communication was English, and the instructor monitored the activities. The students entered this site to complete the required steps for homework completion. These included information posts and a minimum number of discussion comments. In addition to this main channel, the students utilized a digital backchannel in the form of SNS, email, voice calls, and non-digital face-to-face (f2f) communication. The instructor did not monitor these backchannels because they were initially private to the students, so there were few limitations on information type and source. The central box indicates the students and non-students that took part in these communication channels. While the students were asked to only work within their assigned group, the students could get the assistance of non-students from different groups, different classes, and even different schools. All of this was possible because of the digital backchannels and the freedom they provided.

Figure 1

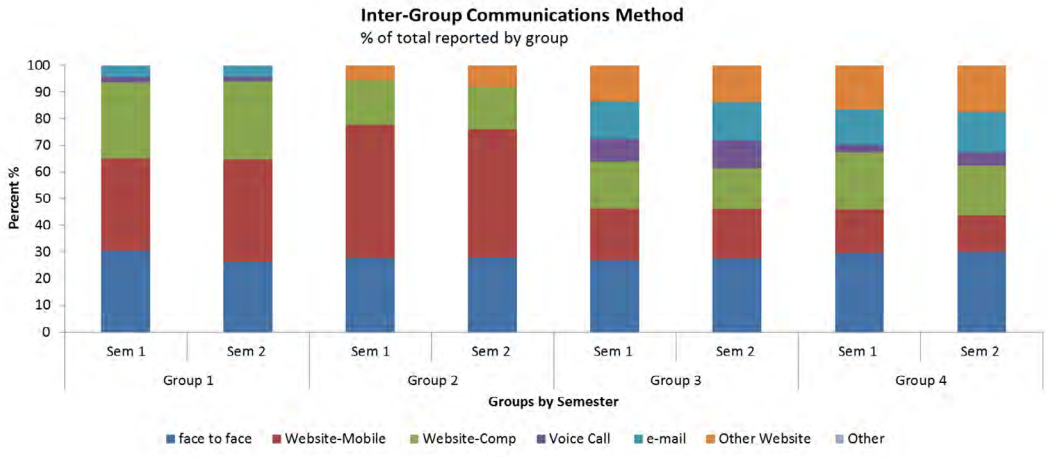
Digital Communication Channels



In Figure 2, the graph presents students' reported methods for inter-group communications coded from their weekly e-journal reports and interviews. Each week, the students were asked to comment on the activities and their group work. This was in addition to pre- and post-interviews. The reported method of within-group communication included f2f, Moodle website via a mobile device, Moodle website via desktop computer, voice phone call, email, Other websites meaning SNS, and any other means of communication- no other means were reported.

Figure 2

Group Communication Methods



F2F was the most common form of communication and was commonly reported to take place between classes or during lunch. Mobile access to the Moodle website was also very popular and included access for reading and posting of information from their group.

For groups two to four, *Other Website* was common and were reported as including SNS sites such as Twitter, Facebook, Line, and Mixi. Also, the other digital backchannels of voice calls and email were frequently reported.

Discussion

The first finding related to how the blurring of boundaries between friends and others increased the development of new peer connections and relationships. The mobile phone use for homework allowed the students to communicate more with their group members because the phone was always present and reduced the need to meet face-to-face. While the students were happy to discuss the topic face-to-face with friends, they did not like the idea of meeting face-to-face with group members that they did not already know. They mentioned that they would not talk with some of them at all if the homework were just face-to-face. However, the same students felt that the mobile phone offered a comfortable backchannel to communicate with those students that they would otherwise not speak with, resulting in an increased chance of making new connections.

Post-Interview Group 2 Student

I: Why are you closer second semester than first? Why did you become close?

S: I think after doing the homework for many times, we exchange the idea many times, we can get familiar with other people. Although in the class, of course, we cannot meet everyone, but from the comment I can feel some characters of them.

I: So you get better friends with them.

S: Yes.

This online method of communication may have lowered barriers for shy and introverted students by providing a more comfortable mode of communication (Camiel et al., 2014; Krishnan & Poleon, 2013). This observed tendency of students to enlarge their online peer networks aligns with the research showing that when students have the freedom to direct discussion in a collaborative environment, they will seek methods relevant to their learning purpose (Yardi, 2006). Mobile SNS that the students previously used only for very private friends started to be used as a backchannel channel for communication with these new peer connections. This may also lead to a reduction in ingroup versus outgroup conflicts (McCarthy et al., 2004; McCarthy & Boyd, 2005). Now, these requests were also seen by friends who were not in their group or even in the same school so that they could get feedback and advice. This aligns with previous observations that these digital backchannels provided more information on the topic of discussion than would be available without them through the main-channel, which may improve critical thinking (Rathakrishnan et al., 2017), and the building of new ideas (Brooke, 1987; Chen & Looi, 2007). These channels were outside the main channel of the Moodle site, even though the Moodle site offered most of the same affordances for communication.

The second finding concerned the students' mobile social networks Twitter and Facebook, which were almost exclusively accessed through their mobile phones. At the beginning of the research, the mobile SNS were firmly in the students' private space for entertainment, so students strongly disliked using them to assist with homework.

Pre-Interviews Group 4 Student

I: So does your group use Twitter to talk about my homework?

S: No. Actually, we don't talk about the homework…

I: Why do you think that is?

S: Mm, I haven't thought that idea [laughs], so I have no idea why that – we think about its private space, so maybe your website is homework/study space, so we divide into space; study or private.

I: Okay. Do you think most students do that?

S: Mm [laughs], maybe, I think.

I: What's the difference between study space and private space?

S: We only do homework things in your website. If we can want other communication – so the space is – can be the – like social communication space.

Post-Interviews Group 4 Student

I: Do you use it [SNS] for homework a little?

S: A little.

I: When?

S: When [pause] we are near deadline of homework.

I: Why?

S: I want to check my group member did it.

I: You don't check my website?

S: I did it – I do it sometimes but Twitter is more [pause] useful to check it because Twitter is [pause] – many people see Twitter many times.

Above is an excerpt from the interview transcript of the same student. The first example is from the pre-interview, and the second example is from the post-interview. These examples show the apparent change in attitude concerning the private and public space in which their mobile social networks belong. They explain clearly in interview one that the

SNS that they access only with their mobile are separated from their school life. When asked why they separate the two, they explain that one is for their private things, and the other is their study space. Here they are identifying a *homework space* and a separate *private space*. This example suggests that most students see the homework website, what they call *your website*, for everyone to use, and mobile social networks for those people with whom they socialize. However, when asked about these same websites in interview two, they say that their group did start to use them for the homework when they were close to the deadline. They were, for the most part, not using the sites for discussion, but instead to post status updates on SNS because they knew that all their friends were continually monitoring these sites with mobile phones, so it was the best source of real-time information updates on the activities of other groups members. This suggests that the students were self-motivated and actively trying to remove any barriers that exist by incorporating these digital backchannels to overcome the limitations that the students felt existed with the main man-channel of information, the Moodle site.

Several members from each group all echoed these feelings about their mobile social networks being personal and separate from homework. One student from group 4 explained in interview one that even though these social networks are open to public viewing, they hide their meaning so only their friends can understand the meaning. At the beginning of the course, mobile SNS, such as Twitter, were identified by the students as something they would not like to use for homework. Some students strongly expressed a need to keep some part of their lives private from school so that they would have a way to relax. Other students identified the SNS as a private place in which homework is not allowed.

However, this attitude changed over the year as the students started to incorporate their mobile social networks into the collaboration process. The homework website did have a mobile notification system that emailed participants after a message post, but the students were allowed at any time to turn this off. Some started to use their private mobile social networks to help keep track of when their group members posted a homework related message. Students reported that group members started to use mobile social networks to let each other know that there were new comments on the homework website. This is similar to the creation of a virtual space for real-time chat communication reported during conferences (McCarthy & Boyd, 2005). As noted in the research (Toledo & Peters, 2010), this change in attitude suggests students will evolve practices that are less disruptive to the everyday technology use patterns to which students have become accustomed.

They used these systems as secondary notification channels by which their friends in the group could request an immediate and private comment on their posted homework message, which has been shown to increase engagement in language classes (Harunasari & Halim, 2019). This desire for more notification of any updates to their group comments suggests that these third-party sites were increasing the quantity and quality of the collaborative experiences by reducing the delay between responses allowing more time to think (Toledo & Peters, 2010). This extra time has been shown to translate into more thoughtful responses from students (Birch & Volkov, 2007; Branon & Essex, 2001; Hanson-Smith, 1997; Kitade, 2008; Ortega, 1997).

The evidence in this section suggests that homework communication on mobile devices leads to these previously private mobile communication paths being exploited for homework.

SNS seem to be providing an alternative digital backchannel information stream, which is a mixing of public homework and private social spaces through the mobile device. Students attempted to make communication more efficient, increasing collaborative quality, which should positively affect the richness of learning.

The third finding was that the students placed their first language (Japanese) into their private communication world while their second language (English) was used for public communications.

Pre-Interviews Group 3 Student

I: Oh okay. Good. When you talk about or chat about homework on Mixi, why don't you do the same thing on my website?

S: I think it's homework, so [Seikakuna bunsyou] correct sentences. I should do [Shinken ni Majime ni Yarubeki…] I should do seriously. So I chat long time on Mixi, but [that] I can't say on the homework page.

I: Okay. Why can't you?

S: Mixi is Japanese but homework is English, so I don't have vocabulary in English, so Japanese is easy, maybe I think.

I: Alright, so you don't use English on Mixi?

S: Yes.

When communicating f2f or using a digital backchannel, the language of communication was usually Japanese. Alternatively, the homework website, which everyone can see, including the course instructor, is an all English language environment. They are allowed short Japanese examples of a few words when explaining a translation, but otherwise, everything must be in English. This public forum for their language was difficult for those students who lacked confidence in their English ability. So, to reduce the chances of a potentially embarrassing public mistake, they used their private mobile communication channels to check their ideas with friends and others. As has been shown, students will seek more comfortable and less threatening forms of communication (Camiel et al., 2014; Carpenter, 2015). This may support students who feel less competent when expressing their opinions to peers (Bry & Pohl, 2017; McNely, 2009), but without the benefits of target language practice.

Conclusion

The study design was a case study adopted for one academic year to gain a deeper understanding of the processes and outcomes of completing collaborative learning activities through mobile devices. Japanese university undergraduate EFL students studying translation formed four case study groups with between five to eight members. Student online website log data, weekly e-journal reports, and pre- and post-study face-to-face interviews

comprised the principal sources of data. The data analysis consisted of content analysis, where data was coded then categorized into themes through an inductive form of thematic analysis coding.

The boundary between friends and others was blurred, so increasing the development of new connections and relationships. The mobile phone use for homework allowed the students to communicate more with their group members because the phone is always present and acting as an always-on digital backchannel. In addition, it reduced barriers and the need to meet f2f with peers they were not familiar with so increasing the peer network (Camiel et al., 2014; Krishnan & Poleon, 2013).

There was an apparent change in attitude concerning the private and public space in which mobile social networks belong. SNS sites firmly in the students' private space moved into the public space by the end of the study to be used as a digital backchannel for updates, answer checking, and notifications from students. This supports the idea that students will evolve practices that are least disruptive to themselves (Toledo & Peters, 2010).

The students placed their first language (Japanese) into their private communication world while their second language (English) was used for public communications. This meant that they fulfilled the requirements of communicating in L2 when the instructor could see, but the private SNS sites allowed a digital backchannel in L1 where they could gather information and ask for advice confidently from group members and others not related to the course or school. This shows that students are using these backchannels as less threatening forms of communication (Camiel et al., 2014; Carpenter, 2015) in which they feel more confident (Bry & Pohl, 2017; McNely, 2009).

This research was limited to Japanese participants at the tertiary level of education. Also, the private nature of the students' SNS service use did not allow for data collection to see the types of information they passed through these digital back channels. Future research would include a more diverse sampling in terms of age, cultural background, and study area. If these findings could be identified across a diverse sample, it would indicate common affordances of the digital backchannels that could inform curriculum design. The inclusion of a digital backchannel system that the researcher could directly observe would allow for the categorization of information exchange types. This would provide a much greater understanding of how, when, and for what purposes the students utilize these backchannels during learning.

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.

Althaus, S.L. (1997). Computer-mediated communication in the university classroom: An experiment with on-line discussions. *Communication Education, 46*(3), 158–174.

Beatty, I.D., Gerace, W.J., Leonard, W.J., & Dufresne, R.J. (2006). Designing effective questions for classroom response system teaching. *American Journal of Physics, 74*(1), 31–39.

Birch, D., & Volkov, M. (2007). Assessment of online reflections: Engaging English second language (ESL) students. *Australasian Journal of Educational Technology, 23*(3), 291–306.
Branon, R.F., & Essex, C. (2001). Synchronous and asynchronous communication tools in distance education. *TechTrends, 45*(1), 36–42.
Brooke, R. (1987). Underlife and writing instruction. *College Composition and Communication, 38*(2), 141–153.
Bry, F., & Pohl, A. Y.-S. (2017). Large class teaching with Backstage. *Journal of Applied Research in Higher Education, 9*(1), 105–128. **https://doi.org/10.1108/JARHE-06-2015-0042**
Caldwell, J.E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE – Life Sciences Education, 6*(1), 9–20. **https://doi.org/10.1187/cbe.06-12-0205**
Camiel, L.D., Goldman-Levine, J.D., Kostka-Rokosz, M.D., & McCloskey, W.W. (2014). Twitter as an in-class backchannel tool in a large required pharmacy course. *American Journal of Pharmaceutical Education, 78*(3), 67–67. **https://doi.org/10.5688/ajpe78367**
Carpenter, J.P. (2015). Digital backchannels: Giving every student a voice. *Educational Leadership, 72*(8), 54–58.
Chen, W., & Looi, C.K. (2007). Incorporating online discussion in face to face classroom learning: A new blended learning approach. *Australasian Journal of Educational Technology, 23*(3), 307–326.
Clark, H.H., & Brennan, S.E. (1991). Grounding in communication. In L.B. Resnick, R.M. Levine, & S.D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 222–233). American Psychological Association.
Cogdill, S., Fanderclai, T.L., Kilborn, J., & Williams, M.G. (2001, January 6). Backchannel: whispering in digital conversation. *Proceedings of the 34th Annual Hawaii International Conference on System Sciences*, Hawai'i, USA.
Donnelly, H.J. (2016). *Exploring Student Interaction and Reflection Through the Use of Digital Backchannel Discussions* (Publication Number 612) Illinois State University. **http://ir.library.illinoisstate.edu/etd/612**
Erickson, T., & Kellogg, W. A. (2000). Social translucence: an approach to designing systems that support social processes. *ACM Transactions on Computer-Human Interaction (TOCHI), 7*(1), 59–83.
Ezzy, D. (2002). *Qualitative analysis: Practice and innovation*. Routledge.
Fies, C., & Marshall, J. (2006). Classroom response systems: A review of the literature. *Journal of Science Education and Technology, 15*(1), 101–109.
Hanson-Smith, E. (1997). *Technology in the Classroom: Practice and Promise in the 21st Century*. Teachers of English to Speakers of the Other Languages.
Harry, D., Green, J., & Donath, J. (2009, April 4–9). Backchan.nl: Integrating backchannels in physical space. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Boston, MA, USA.

Harunasari, S.Y., & Halim, N. (2019). Digital Backchannel: Promoting Students' Engagement in EFL Large Class. *International Journal of Emerging Technologies in Learning (iJET), 14*(07), 163–178.

Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. *Journal of Computing in Higher Education, 15*(1), 46–64.

Hennessy, C.M., Kirkpatrick, E., Smith, C.F., & Border, S. (2016). Social media and anatomy education: Using twitter to enhance the student learning experience in anatomy. *Anatomical Sciences Education, 9*(6), 505–515.

Honeycutt, C., & Herring, S.C. (2009). Beyond microblogging: Conversation and collaboration via twitter. *Proceedings of the Forty-Second Hawai'i International Conference on System Sciences* (HICSS-42), Los Alamitos, CA.

Jarrett, K., & Devine, M.A. (2010). How to use backchanneling in your classroom. *The Education Digest, 76*(1), 41–44.

Kay, R.H., & LeSage, A. (2009). Examining the benefits and challenges of using audience response systems: A review of the literature. *Computers & Education, 53*(3), 819–827.

Kitade, K. (2008). The role of offline metalanguage talk in asynchronous computer-mediated communication. *Language Learning & Technology, 12*(1), 64–84.

Krishnan, J., & Poleon, E. (2013). Digital backchanneling: A strategy for maximizing engagement during a performance-based lesson on Shakespeare's Macbeth. *Teaching English with Technology, 13*(4), 38–48.

Lantz, M.E. (2010). The use of 'clickers' in the classroom: Teaching innovation or merely an amusing novelty? *Computers in Human Behavior, 26*(4), 556–561.

McCarthy, J.F., Boyd, D., Churchill, E.F., Griswold, W.G., Lawley, E., & Zaner, M. (2004, November 6–10). *Digital backchannels in shared physical spaces: attention, intention and contention* [Paper presentation]. 2004 ACM Conference on Computer Supported Cooperative Work, Chicago, Illinois, USA.

McCarthy, J.F., & Boyd, D.M. (2005). *Digital backchannels in shared physical spaces: experiences at an academic conference* [Paper presentation]. CHI'05, Portland, OR, USA.

McNely, B. (2009). Backchannel persistence and collaborative meaning-making. *Proceedings of the 27th ACM international conference on Design of communication*. Bloomington, Indiana, USA.

Ortega, L. (1997). Processes and outcomes in networked classroom interaction: Defining the research agenda for L2 computer-assisted classroom discussion. *Language Learning & Technology, 1*(1), 82–93.

Rathakrishnan, M., Ahmad, R., & Suan, C.L. (2017, April 3–5). Online discussion: Enhancing students' critical thinking skills. *AIP Conference Proceedings*. The 2nd International Conference on Applied Science and Technology 2017 (ICAST'17), Kedah, Malaysia.

Rogoff, B., Paradise, R., Arauz, R.M., Correa-Chávez, M., & Angelillo, C. (2004). Firsthand learning through intent participation. *Análise Psicológica, 22*(1), 11–31.

Ross, C., Terras, M., Warwick, C., & Welsh, A. (2011). Enabled backchannel: Conference Twitter use by digital humanists. *Journal of Documentation, 67*(2), 214–237.
Schick, R.J. (2005). *A study of student conversation in text and audio during webcast lectures*. University of Toronto.
Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction, 4*(4), 295–312.
Toledo, C., & Peters, S. (2010). Educators' perceptions of uses, constraints, and successful practices of backchanneling. *in education, 16*(1), 75–88.
White, S. (1989). Backchannels across cultures: A Study of Americans and Japanese. *Language in Society, 18*(1), 59–76. **http://www.jstor.org/stable/4168001**
Williams, K.C., Marie. (2000). Reproduced and emergent genres of communication on the World Wide Web. *The Information Society, 16*(3), 201–215.
Wolf, J.P. (2008). The effects of backchannels on fluency in L2 oral task production. *System, 36*(2), 279–294.
Yankelovich, N., McGinn, J., Wessler, M., Kaplan, J., Provino, J., & Fox, H. (2005). Private communications in public meetings. CHI'05, Portland, OR, USA.
Yardi, S. (2006). The role of the backchannel in collaborative learning environments. *ICLS '06 Proceedings of the 7th international conference on Learning Sciences*. Bloomington, Indiana, USA.
Yardi, S. (2008). *Whispers in the classroom*. The MIT Press. **https://doi.org/10.1162/dmal.9780262633598.143**
Yates, K., Birks, M., Woods, C., & Hitchins, M. (2015). # Learning: The use of back channel technology in multi-campus nursing education. *Nurse Education Today, 35*(9), e65–e69.
Yngve, V. H. (1970). On Getting a Word in Edgewise. *Papers from the Sixth Regional Meeting of Chicago Linguistic Society*, Chicago, IL, USA. **https://ci.nii.ac.jp/naid/10026753338/en/**

Author's Bio

Dr. Peter Ilic currently holds the position of Associate Professor in the Center for Language Research at the University of Aizu, in Japan. Information and Communications Technology in Education remains the principal focus of his ongoing research which has resulted in several peer-reviewed publications, as well as contributed presentations.

https://doi.org/10.37546/JALTSIG.CALL2020.6

6
Comparing Writing vs. Smartphone Tapping Speed

Bradford J. Lee, Fukui University of Technology

Abstract

As COVID-19 pushes the world into emergency remote teaching mode, many teachers struggle with designing online or computer-mediated language learning activities due to having little to no prior experience. It is imperative to understand how the use of technology affects not only the processes of learning, but also the outcomes. Particularly in Japan, where smartphone ownership among adults aged 18–34 topped 96% in 2018 (Pew Research Center, 2019), many students use their phones to access their schools' learning management systems (LMS) and complete assignments (i.e., mobile learning). The current study therefore sought to elucidate how different writing media can affect the execution of a simple writing task by examining the differences in transcription speeds between handwriting and 'tapping' on a smartphone. A total of 176 participants were divided into 3 groups (L1-English, writing in English; L1-Japanese, writing in Japanese; and L1-Japanese, writing in English), and their times-on-task were recorded. While no difference was found for the L1-English group, the L1-Japanese groups were found to be significantly faster at one task than the other (tapping in Japanese and writing in English). Pedagogical implications suggest the need for instructors to be aware of the extreme difficulty language learners may have when using mobile devices for writing tasks.

COVID-19が世界を緊急リモートティーチングモードに移行させるにつれ、多くの教師は、オンラインまたはコンピューターを介した言語学習活動の設計においてほとんどまたはまったく経験がないことに苦労しています。テクノロジーの使用が学習のプロセスだけでなく結果にもどのように影響するかを理解することが不可欠です。特に日本では、18歳から34歳の成人のスマートフォンの所有率が2018年に96%を超えたため（Pew Research Center、2019）、多くの生徒がスマートフォンを使用して、学校の学習管理システムや課題（モバイル学習など）にアクセスしています。したがって、現在の研究では、スマートフォンでの「タップ」と手書きの転写速度の違いを調べることにより、さまざまな書き込みメディアが単純な書き込みタスクの実行にどのように影響するかを解明しようとしました。合計176人の参加者を3つのグループ（L1英語、英語で書く、L1日本語、日本語で書く、およびL1日本語、英語で書く）に分けて、作業時間を記録しました。L1-Englishグループの違いは見つかりませんでしたが、L1-Japaneseグループは片っ方が遥かに速かったことがわかりました（日本語ではタップ、英語では手書き）。教育的影響としては、講師がモバイルデバイスを使用してタスクを作成する際に、言語学習者が抱える可能性のある非常に困難に気づく必要があることが挙げられます。

Keywords: **CALL, MALL, m-learning, smartphones, handwriting, tapping, writing tasks**

Mobile learning (m-learning) has represented the next stage in the evolution of computer-mediated language learning, with most apps, websites, and most importantly, university learning management systems (LMS) now having interfaces for both forms of technology. This feature is specifically advertised by LMS providers as offering enhanced convenience for both teachers and students, ostensibly lowering the barrier for classrooms to transition online. This is no doubt in response to trends in global smartphone usage, represented in one survey which reported that 97.9% of users at one university (students and faculty/staff) spent more than 3.5 hours daily on their phones (Berolo et al., 2011). Particularly in Japan, where smartphone ownership among adults aged 18–34 surpassed 96% in 2018 (Pew Research Center, 2019), university students typically have portable Internet-capable devices (e.g., smartphones, tablets) with them in class at all times (e.g., 99.56%; Lee, 2019a). Precursor research, which asked university students if they would choose smartphone- over paper-based submission for English classwork if given the choice, found that 75.56% either said yes, or that they would have no issue against doing so (Lee, 2019a). When asked follow-up questions, such as if they thought using a smartphone would affect their written English production, 40% reported that they didn't think it would make a difference, along with 25.33% who thought that they would probably write *more* if they could compose on their phones (Lee, 2019a).

While the example provided above reflects the sentiment of students given a hypothetical situation, the actual reality of using smartphones for composition in an educational setting has largely been unexplored. In fact, those very same students who reported believing that composing on their smartphone would either make no difference, or allow them to write more, actually produced significantly *less* than their handwriting peers did (Lee, 2019b, 2019c). It should therefore be noted that a high degree of familiarity with smartphone technology should not preclude a careful examination of the cognitive intricacies when using it in a new way (i.e., for compositions of longer length), especially in the case of one's second language (L2). This study therefore set out to answer the following research questions:

1. How does transcription speed differ between handwriting and smartphone tapping for L1 speakers of English and Japanese, writing in their native languages?
2. How does transcription speed differ between handwriting and smartphone tapping for L1 speakers of Japanese, writing in an L2 (e.g., English)?

Review of the Literature

The Importance of Writing Speed/Fluency

Writing is the conversion of ideas and language into a tangible form that can be viewed on a page. Text generation can be conceptualized as necessitating four components: long-term memory (accessed while composing), short-term memory (activated while reviewing), executive functions (e.g., making use of strategies, revising), and the act of transcription itself (e.g., handwriting, keyboarding, or tapping) (Berninger et al., 2002). Of these four, transcription ability is the fundamental skill which needs to be developed, as without it, the other components are impossible. Once mastery of fluent and accurate handwriting is achieved (in addition to competence in spelling), then executive processes for composition

writing can develop, which gradually transition to become more self-regulated (Berninger, 1999).

It has also been suggested that these components of writing all draw from the same limited cognitive resources of the writer. Essentially, increased demands by one skill (e.g., transcription), will cause other processes to suffer (i.e., executive control, planning) (Torrance & Galbraith, 2005). This has been demonstrated empirically in several ways. Studies of younger children (i.e., under age 11) have shown that they produce better essays when they dictate them, rather than write them (De La Paz & Graham, 1995), though this trend reverses past this age, as handwriting skills become more automated (Bereiter & Scardamalia, 1987). However, handwriting speed is still linked with overall exam quality into adulthood, with slower handwriting being correlated to lower performance on exam essays (Connelly et al., 2005).

Handwriting vs. Digital Writing

Manual and digital writing have several major differences, both physical and cognitive. Physically, writing by hand makes use of implements such as pencils, which are held in the hand and manipulated using fine-motor control to form letters on the page. Visual sensory feedback, as the writer monitors the written production, is integrated with kinesthetic feedback the writer receives through the tip of the pen to create accurate pressure, timing, and angles of movement (e.g., Mangen & Velay, 2010). Formation of the letters themselves evokes long-term memory as to their standard shape, combined with the writer's own stylistic qualities which make each person's handwriting unique. In addition to penmanship, other paralinguistic features are easily infused into a person's writing (sometimes unconsciously) which can indicate changes in emotive state (e.g., leisurely or hurried writing) or the desire to communicate intimacy or group membership, such as *marumoji* (rounded characters), a style of handwriting developed and used exclusively among young women and schoolgirls in Japan in the 1970s (see Hansen, 2015; Kataoka, 1997; Satake, 1980).

Compared to the nuanced and emotive dimensions of handwriting, digital text entry is merely an exercise in simple memory. There is no graphomotor component as typing/tapping requires only the press of a button; the only difference between irrespective letters is their physical locations on the keyboard and the fingers used to press them. This involves much less muscle activity than writing and does not require fine motor control (Ko et al., 2015). Haptic feedback varies depending on the choice of keyboard and disappears altogether in the case of smartphones which have virtual keypads. (Note that smartphones typically make electronic keyboard clicks to simulate haptics, but these can be eliminated by turning the phone to silent/manner mode.) One key difference with regard to keyboard typing is the separation of the visual and motor spaces, that is, a fluent typist typically watches the letters appear on the monitor while the hands/fingers operate the keyboard out of view. This spatio-temporal separation contrasts with handwriting, where the visual and motor spaces are unified on the tip of the writing implement. In the case of smartphone users, the visual and motor spaces are much closer together, though most users cannot blind type and therefore must look at their fingers while tapping. Typing posture also varies significantly

between users, for example, one- vs. two-finger, thumb vs. index finger, right- vs. left-handed, and so forth. (Palin et al., 2019).

In Japan, a special smartphone keyboard has been developed, resembling a 10-key keypad (see Figure 1). Each key represents the top consonant phoneme (i.e., /~a/), with the other phonotactically possible variations (i.e., /~i/, /~u/, /~e/, and /~o/ appearing as a submenu in a cross-pattern (see Figure 2), making selection possible with either a tapping or a swiping motion.

Figure 1

Japanese Smartphone (iPhone) Keyboard

Figure 2

Japanese iPhone Keyboard with Phonotactic Variations Displayed

Unfortunately, while the main topic of this study is smartphone tapping speed, very little work has been done in this area. At the time of this writing, the study by Palin et al. (2019) is the only known large-scale investigation into smartphone tapping proficiency. In their study, they reported an average tapping speed of 36.2 WPM (words per minute) with 2.3% uncorrected errors. However, L1 speakers of English were much faster than speakers of other languages who did not typically type in English (M = 37.8 vs. 25.6, respectively). In

addition, while tapping style impacted performance, keyboard typing skills (i.e., touch-typing ability) did not. Support for these findings can be found from an ergonomic perspective, as two-handed typing at chest level has been shown to activate the least muscle activity in the trapezius (the muscles which connect the shoulders and neck) and the fastest typing speed, with the slowest being single-hand thumb tapping (Ko et al., 2015). Interestingly, Palin et al., (2019) suggested that intelligent text entry methods of smartphones affected performance in various ways; autocorrection positively affected speed while word prediction actually hindered it.

Study Design

Participants

The main site for recruitment of the L1 Japanese speakers in the current study was the self-access language learning center of a small, private Japanese university which specializes in engineering and industrial sciences. A total of 144 volunteers agreed to take part, comprised of first-year (n = 39), second-year (n = 61), and third-year (n = 44) students (i.e., 18–21 years of age). The rural location and low number of L1 English residents in the area make for an English as a foreign language (EFL) condition, as students at the university have very few opportunities to use English outside of class. In addition, as this university does not have an English degree program, all participants were majoring in other disciplines and only enroll in English courses to satisfy their basic requirements for graduation. All participants were Japanese nationals with no history of living in any other country.

Each year, students are given an abridged TOEIC® (Test of English for International Communication) Bridge test (only the listening and reading sections), the scores of which are used for class placement. The four class divisions, and their approximate level according to the Common European Framework of Reference for Languages (CEFR) are as follows: High (CEFR-B1 and higher), Medium (CEFR-A2 to B1), Low (CEFR-A2), and Basic (CEFR-A1 and lower). While each participant's English proficiency was not assessed as part of this study, their class affiliations allow for the assumption that they are generally between CEFR-A2 and B1 (High, n = 3; Medium, n = 76; Low, n = 53; Basic, n = 12), though levels are an aggregate of their listening and reading abilities.

The L1 English speakers were recruited electronically through postings on social media sites for English-speaking expatriates living in the same area as the university. All 32 respondents were EFL instructors (age: M = 28.4) who had come to Japan from inner circle countries (see Kachru, 1990) as adults.

Materials and Methods

Currently, among researchers of typing speed/proficiency, there is no standard consensus regarding exactly how measurements should be taken. For example, some researchers count the number of words produced in the first two minutes of writing (e.g., Phelps et al., 1985; Wallen et al., 1996), or the number of times the alphabet can be written in lower- and upper-case in a one-minute span (e.g., Berninger et al., 1991). However, as this study sought

to simulate students transcribing a text shown to them in class (akin to notetaking), it was decided that the time would be measured by how long it took for students to reproduce the entire text.

Two original passages were created, one in English (Appendix A) and the other in Japanese (Appendix B). Care was taken in the design of the English passage to ensure that: a) it contained all 26 letters of the alphabet, b) it contained both capital and lowercase, c) it contained numerals, d) it contained several types of punctuation, and e) it was graded at Level 3 of the New General Service List (Browne et al., 2013). These considerations were made to account for the fact that smartphone keyboards hold items such as punctuation, numerals, and uppercase letters on different screens, necessitating extra keystrokes for their input that proficient typists must master. In addition, while the study was one of transcription, not comprehension, it was believed that vocabulary items that were too far outside of the participants' English level may have affected transcription speeds if students struggled to hold new words in their working memories. The use of a transcription exercise versus a free writing one has the advantage of eliminating thinking time which would possibly confound the results (see also Horne et al., 2011).

The English passage was 164 words long, with a direct translation serving as the Japanese passage. The Japanese version consisted of 337 characters, which included all three forms of writing: logographic kanji and syllabic hiragana/katakana. All of the punctuation and numerals were carried over directly to mirror the English version.

Participants from each group were given their respective reading passages on a sheet of A4-sized paper and told that they would be asked to time how long it took for them to transcribe them as quickly and as accurately as possible. The instructions were explicit that the text was to be transcribed exactly as is, including punctuation, spacing, and capitalizations, and if any mistakes were made during the process, they were required to correct them immediately. However, as it was a timed exercise, they were encouraged to proceed as quickly as possible and stop their timers immediately upon completion. Once manual transcription tasks were completed, the smartphone input was collected by way of an online Google Doc which participants opened by scanning a QR code. The instructions were the same as the writing task; all typos had to be immediately corrected and final times were recorded directly onto the Google Doc and submitted along with the tapped text. The order of data collection was counterbalanced so that half of the participants from each group did the smartphone entry first, followed by the handwriting task. All submissions were post-evaluated to ensure accuracy of reproduction; any errors detected were accounted for by adding one second to the reported time for that submission.

Results

Overall Results

All times were converted into seconds and rounded to two decimal places. These figures were then input into a dataset for IBM's Statistical Package for the Social Sciences (SPSS), Version 24. All statistical calculations were performed using this software. Descriptive statistics (*M* and *SD*) were calculated, as reported below (Table 1).

Table 1

Descriptive Statistics of Transcription Times (in seconds)

	Handwriting		Tapping	
Group	***M***	***SD***	***M***	***SD***
L1-E (n = 32)	333.31	58.34	367.62	131.12
L1-J (n = 74)	419.70	99.29	351.05	108.63
L2-JE (n = 70)	492.26	119.83	1104.40	344.38

Note. L1-E: L1 speakers of English; L1-J: L1 speakers of Japanese; L2-JE: L1 speakers of Japanese, transcribing English

As can be seen by the times reported in Table 1, groups L1-E and L2-JE were faster at transcribing their passages by hand (333.31 vs. 367.62 and 492.26 vs. 1104.40, respectively), while group L1-J showed faster times when tapping on their smartphones (419.70 vs. 351.05). Shapiro-Wilk tests for normality revealed that all groups were non-normally distributed (see Table 2).

Table 2

Shapiro-Wilk Normality of Distribution Output

	Handwriting			Tapping		
Group	**Statistic**	***df***	***p***	**Statistic**	***df***	***p***
L1-E	.93*	32	.05	.92*	32	.03
L1-J	.80*	74	< .001	.96*	74	.02
L2-JE	.94*	70	.002	.90*	70	< .001

Note. * indicates rejection of the null hypothesis, i.e., data is non-normally distributed

Despite the results shown in Table 2 indicating that the data were all in a non-normal distribution, it was decided that paired-samples *t*-tests were suitable to compare the differences in means between the handwriting and tapping speeds for each group. Traditionally, parametric tests such as *t*-tests and analysis of variance (ANOVA) assume a normal distribution. However, it has been suggested that these tests are robust enough to be used with non-normal data, especially in the case of large sample sizes (i.e., over n = 50), as was the case with the two experimental groups involving Japanese students (see Du et al., 2017; Lix et al., 1996). Statistical significance of the differences in means was therefore assessed by paired-samples *t*-tests, with effect sizes expressed as Cohen's *d* values, shown in Table 3.

Table 3

Output of t-Tests (Comparing Handwriting to Tapping Speeds)

	T	*df*	*p*	*d* [95% *CI*]
L1-E	−1.74	31	.09	.30 [−.06, .65]
L1-J	4.27**	73	< .001	.66 [.32, 1.00]
L2-JE	−15.87**	69	< .001	2.16 [1.67, 2.66]

** denotes significance at the p < .001 level

As can be seen in the output of the *t*-tests in Table 3, the difference between the means for group L1-E was non-significant (p = .09), resulting in an effect size of d = .30 with 95% confidence intervals (*CI*) falling on both sides of zero. Taken together, this strongly indicates that statistically, there was no difference between the mean times native speakers of English take to write by hand versus by smartphone.

The results for groups L1-J and L2-JE did, however, result in statistically significant differences between means. Both groups had an extremely small p-value (< .001) and effect sizes with both 95% *CI*s above zero. From these results, it can be concluded that members of group L1-J were able to transcribe the Japanese text significantly faster by smartphone to a medium effect size of d = .66. Conversely, members of L2-JE transcribed the English text significantly faster manually to an extremely large effect size of d = 2.16 (for L2-research field specific interpretations of effect sizes, see Plonsky & Oswald, 2014).

Discussion

The results presented in the previous section were significant for several reasons. First, the finding that transcription speeds of English among L1 speakers were not significantly different between paper and smartphone media was informative and contributes to the scant literature examining smartphone tapping speeds. While there can be no doubt that fluent typists (i.e., on a typewriter/computer) can type faster than a person can write, it seems that this advantage is severely reduced, or disappears entirely, when using a smartphone's keyboard, falling in line with Palin et al.'s (2019) findings. It is reasonable to assume that the smaller size of a smartphone, which reduces the use of all 10 fingers to just one or two, would result in a reduction of typing fluency, though it was surprising that the reduction was so great as to achieve parity with handwriting. However, it should be mentioned that even at the relatively short length of the sample passage (i.e., 164 words for English; 337 characters for Japanese), fatigue of the hand and fingers became an issue as some participants could be observed rubbing or shaking their hands while writing. As this sort of fine-motor control is not utilized in tapping/typing, it is probable that average handwriting speeds would decline over a longer span while tapping speeds would not; future research is needed to confirm this, or at what point a decline in speed begins to occur.

The result that Japanese L1 speakers can tap in Japanese faster than they can write fell in line with predictions. As Japanese is a logographic language which uses multi-stroked characters, a single graph can often be extremely complex, requiring a high degree of dexterity and fine-motor control to write. In comparison, smartphone input in Japanese is done

phonetically, in combination with a form of word prediction, resulting in a much faster input requiring significantly less motions. For example, the Japanese word, 鬱 (*utsu* = depression) requires 29 separate strokes of a pen, but only three taps (including selection from a prediction list) to enter on a smartphone. While such an extreme case was not present in the current study, several basic words in the text were over 10 strokes each, including 新, 業, 頑, 損, and 暇, which are 13 stroke characters (see Appendix B).

Finally, the finding that Japanese L1 speakers' transcription speed of L2 English is significantly slower on smartphone when compared to handwriting is a novel finding. As the exercise was one of simple transcription, issues such as lexical familiarity and spelling can be ruled out as a source of the significant speed reduction. The fundamental issue is therefore believed to be one of English keyboard familiarity. This finding departs somewhat from Palin et al.'s (2019) suggestion that touch-typing training did not impact tapping speed. However, as the smartphone version of a Japanese keyboard is entirely different from an English one (see Figures 1 & 2), it would be reasonable to assume that Japanese users may face at least some level of difficulty in acclimating to a different keyboard (although computer keyboards in Japan are in the standard QWERTY layout). A multivariate experiment involving handwriting, smartphone tapping, and computer typing could possibly elucidate whether this is the case, or if the issue lies elsewhere.

This result could also possibly explain the results of precursor research that found that Japanese L1 smartphone tappers produced significantly less text than their handwriting counterparts (Lee, 2019b; 2019c). As tapping appears to take significantly more time, cognitive resources can be argued to be thus diverted from the language task towards the mechanical task of tapping. This would reduce the quality/quantity of prose that could be produced, especially if assuming approximately equivalent lengths of time-on-task (precursor research involved free-writing tasks where students could stop writing at any time).

Pedagogical Implications

Pedagogically speaking, this study has several implications for language classrooms. Precursor research revealed that most Japanese students view smartphone use for completing English writing assignments favorably (Lee, 2019a). This view is supported by the current study which demonstrated that in Japanese, tapping on a smartphone is a significantly faster method of text generation. However, the assumption that the same would hold true for English L2 production would be a serious miscalculation on the students' part. From the language teachers' perspective, based on the finding that the writing medium does not significantly affect writing speed for English L1 speakers, it is possible that instructors will not realize that it makes a significant difference for their students. Instructors (including Japanese L1 instructors of English) need to be cognizant of the confounding variable that writing medium introduces into the performance of an L2 writing task.

Limitations and Future Directions

This study had several methodological limitations, many of which have already been detailed. As the outcome measure of the task was time, the creation of the transcription text

had a slight impact on the results. The English text was purposely composed to include features such as capitalizations, numerals, and punctuation; the absence of these features would most likely result in shorter transcription times on the smartphone as they required extra keystrokes. The Japanese text also is slightly dependent on the choice of lexical items, as some Japanese characters are more complicated than others. The inclusion of extremely complex characters would therefore increase the time it takes to manually write them, though this would not affect smartphone input to the same degree. Future studies should experiment with different materials, such as authentic texts, to see if similar patterns can be detected to improve generalizability.

The length of the text, and the way the times were measured are also potential limitations. As mentioned in the Materials and Methods section above, the current study utilized one of many possible methods of determining tapping/writing fluency. As calculations of writing speed vary widely between studies, this somewhat limits generalizations that can be made. However, the materials and methods employed in this case were representative of a typical scenario that students at this university could be expected to face, and therefore meaningful in the current context.

One final consideration is that the participants in the study were aware that they were being timed. They were therefore fully intent on the task and cognizant of the clock, with many writers commenting afterwards that their hands began to cramp, but they persevered on to achieve the best time possible. The results produced here can therefore be thought of as the maximum speed for the respective media. However, as natural production would not take place at such break-neck speeds, it would be informative to measure writing times under more realistic circumstances. While handwriting times would most definitely increase, it would be interesting to see how tapping speeds would be affected.

Conclusion

Smartphone tapping speed is an area that has still yet to be thoroughly investigated. While some studies have sought to measure the various speeds of different tapping styles or body positions (e.g., Ko et al., 2015; Palin et al., 2019), this is the first known study which examined how tapping speed compared to writing speed in participants' L1 versus an L2. To measure writing speeds without confounding variables such as creativity, lexical choice/knowledge, and spelling proficiency, graded passages were created for participants to transcribe, either by hand or by smartphone. The results suggested that there was no difference in speed between the two media for L1 speakers of English. However, L1 speakers of Japanese were found to be significantly faster at tapping by smartphone. It is believed that despite the complexity of the Japanese writing system, the ease of digital entry (using a phonetic system along with a specialized smartphone keyboard layout) contributed to this speed advantage. Most importantly, it was found that this trend was reversed when L1 speakers of Japanese wrote in L2 English. When writing in English, Japanese participants were significantly slower tapping on their phones, to a very large effect size (d = 2.16). This study hopes to elucidate potential problem areas for language instructors to consider while creating lessons and assessments during online learning.

References

Bereiter, C., & Scardamalia, M. (1987). *Psychology of written composition*. Erlbaum.

Berninger, V.W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly, 22*(2), 99–112. **https://doi.org/10.2307%2F1511269**

Berninger, V.W., Mizoikowa, D.T., & Bragg, R. (1991). Scientific practitioner: Theory-based diagnosis and remediation of writing disabilities. *Journal of School Psychology, 29*(1), 57–59. **http://doi.org/10.1016/0022-4405(91)90016-K**

Berninger, V., Vaughan, K., Abbott, R., Begay, K., Byrd, K., Curtin, G., Minnich, J., & Graham, S. (2002). Teaching spelling composition alone and together: Implications for the simple view of writing. *Journal of Educational Psychology, 94*(2), 291–304. **https://doi.org/10.1037/0022-0663.94.2.291**

Berolo, S., Wells, R.P., & Amick, B.C. III. (2011). Musculoskeletal symptoms among mobile hand-held device users and their relationship to device use: A preliminary study in a Canadian university population. *Applied Ergonomics, 42*(2), 371–378. **https://doi.org/10.1016/j.apergo.2010.08.010**

Browne, C., Culligan, B., & Phillips, J. (2013). The New General Service List. **http://www.newgeneralservicelist.org/**

Connelly, V., Dockrell, J., & Barnett, J. (2005). The slow handwriting of undergraduate students constrains overall performance in exam essays. *Educational Psychology, 25*(1), 99–107. **https://doi.org/10.1080/0144341042000294912**

De La Paz, S., & Graham, S. (1995). Dictation: Applications to writing for students with learning disabilities. In T. Scruggs & M. Mastropieri (Eds.), *Advances in learning and behavioral disabilities* (pp. 227–247). JAI Press.

Du, H., Zhang, Z., Yuan, K.H. (2017). Power analysis for t-Test with non-normal data and unequal variances. In L.A. van der Ark, M. Wilberg, S.A. Culpepper, J.A. Douglas, & W.C. Wang (Eds.), *Quantitative psychology* (pp. 373–380). Springer: Springer Proceedings in Mathematics & Statistics: Vol. 196. **https://doi.org/10.1007/978-3-319-56294-0_32**

Hansen, K. (2015). Authenticity in Japanese cell phone novel discourse. *US–Japan Women's Journal, 48*(2015), 60–78. **https://doi.org/10.1353/jwj.2015.0004**

Horne, J., Ferrier, J., Singleton, C., & Read, C. (2011). Computerized assessment of handwriting and typing speed. *Educational & Child Psychology, 28*(2), 52–66.

Kachru, B. (1990). World Englishes and applied linguistics. *World Englishes, 9*(1), 3–30. **https://doi.org/10.1111/j.1467-971X.1990.tb00683.x**

Kataoka, K. (1997). Affect and letter writing: Unconventional conventions in casual writing by young Japanese women. *Language in Society, 26*(1), 103–136. **https://doi.org/10.1017/S0047404500019424**

Ko, P-H., Hwang, Y-H., & Liang, H-W. (2015). Influence of smartphone use styles on typing performance and biomechanical exposure. *Ergonomics, 59*(6), 821–828. **https://doi.org/10.1080/00140139.2015.1088075**

Lee, B. (2019a). Japanese tertiary students' access to smartphone and their feelings regarding their use in the EFL classroom. *Memoirs of Fukui University of Technology, 49,* 216–224. **http://hdl.handle.net/10461/29134**

Lee, B. (2019b). A case study of writing task performance: Smartphone input vs. handwriting. *Memoirs of Fukui University of Technology, 49,* 225–231. **http://hdl.handle.net/10461/29135**

Lee, B.J. (2019c). Smartphone tapping vs. handwriting: A comparison of writing medium. *EuroCALL Review, 28*(1), 15–25. **https://doi.org/10.4995/eurocall.2020.12036**

Lix, L.M., Keselman, J.C., & Keselman, H.J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance "F" test. *Review of Educational Research, 66*(4), 579–619. **https://doi.org/10.3102%2F00346543066004579**

Mangen, A., & Velay, J. (2010). Digitizing literacy: Reflections on the haptics of writing. In M.H. Zadeh (Ed.), *Advances in haptics* (pp. 385–401). IntechOpen. **https://doi.org/10.5772/8710**

Palin, K., Feit, A.M., Kim, S., Kristensson, P.O., & Oulasvirta, A. (2019). How do people type on mobile devices? Observations from a study with 37,000 volunteers. In *MobileHCI '19: Proceedings of the 21st international conference on human-computer interaction with mobile devices and services* (pp. 1–12). ACM. **https://doi.org/10.1145/3338286.3340120**

Pew Research Center. (2019). *Spring 2018 Global Attitudes Survey.* [Q.46]. Available from Pew Research Center Website: **https://www.pewresearch.org/global/2019/02/05/smartphone-ownership-is-growing-rapidly-around-the-world-but-not-always-equally/pg_global-technology-use-2018_2019-02-05_0-03/**

Phelps, J., Stempel, L., & Speck, G. (1985). The children's handwriting scale: A new diagnostic tool. *Journal of Educational Research, 79*(1), 46–50. **https://doi.org/10.1080/00220671.1985.10885646**

Plonsky, L., & Oswald, F.L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. **https://doi.org/10.1111/lang.12079**

Satake, H. (1980). Wakamono zasshi no kotoba: Shin-genbun itchi-tai (The language of youth magazines: A new form of *genbun itchi*). *Gengo Seikatsu, 343*, 46–52.

Torrance, M. & Galbraith, D. (2005). The processing demands in writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research*. Guilford Press.

Wallen, M., Bonney, M., & Lennox, L. (1996). *The Handwriting speed test.* Helios.

Appendix A

Text for Transcription (English)

My New Year's Resolutions for 2020

This year, I made 3 New Year's Resolutions. First, I want to exercise more! Last year, I gained about 5kgs because I was lazy and didn't exercise much. I think I should start weight training and go jogging at least 2 times a week. My goal is to lose 10kgs by December 31st.

Next, I will try hard at my job. I want to take a trip with my family to the United States of America in this Fall. My family has 6 members: me, my mother and father, my sister and brother, and my grandmother. I need to save about 1 million yen!

Finally, I want to keep healthy this year. Last year, I caught influenza in the winter, and I couldn't go to school for a full week. I was quite bored staying home alone, and I missed a lot of work. This year I will eat healthy food and exercise to make my body strong.

Appendix B

Text for Transcription (Japanese)

新年の抱負、2020年！

さぁ、今年はどんな1年になるでしょう?まずは健康を大切にしたいと思っています。去年はずっと運動できず5キロほど太ってしまいましたので今年こそはまじまじと運動したいと思います。毎週少なくとも2回ジョギングすることと筋トレを励みたいところです。

次は、仕事も頑張りたいです。今年の夏か秋あたりにアメリカへの家族旅行をしたいなぁ～。家族は6人：自分、父、母、お姉ちゃんと弟、そしておばあちゃん。全員を連れて行くとすると、百万円が必要ですね。

最後に、風邪を引かないように気を付けます！去年の冬にインフルエンザをかかって、一週間ずっと学校に行けなかった。暇過ぎて退屈だったし、その分授業に出られなかったので損しました。今年は健康的な食事と体を動かすことで健康を守ります！

Author's Bio

Dr. Bradford J. Lee, Ed.D. is an associate professor in the Organization of Fundamental Education at Fukui University of Technology (Fukui Kogyo Daigaku) in Fukui City, Japan. His main research interests are phonetics/phonology and pronunciation instruction. Motivation, noticing, and writing media have also been explored.

https://doi.org/10.37546/JALTSIG.CALL2020.7

7

Law Undergraduates' Understanding and Appropriation of Arguments in Online Essay Writing Tutorials

Suman Luhach, Bennett University, India

Abstract

The art of constructing an assertive argument is a crucial lifelong skill for law students to master. The English language teachers at Bennett University Law School introduce the topic of argumentative essay writing to its first-year law undergraduates and teach the basic structure of an argument in a standard five paragraph argumentative essay and then gradually elaborate on the content. The pedagogy makes use of the online platforms of i-Learn learning management and Clarity English programs (customised English language teaching software) to engage students in essay writing tutorials. The study analyses how the students develop the understanding of framing a strong argument and move towards attaining appropriation in it. This was done by comparing pre-test and post-test results of the control and experimental groups and by content analysis of the transcripts. The target group in the study included 120 students of BA LLB (Hons) who are divided into control and experimental groups. The study also tries to figure out the comparative advantage of classroom teaching in physical settings, online group discussions on i-Learn forums, and individual practice sessions on Study Skills Success and the Practical Writing program offered by Clarity English. This was done through student survey analysis. The overall data involved pre-test and post-test essay writing transcripts, online discussion forum transcripts on i-Learn, and student surveys. This educational intervention is an attempt to assess and design the best teaching practice with available resources for teaching argumentative essay writing to law undergraduates.

主張する議論を構築する技術は法科の学生が獲得すべき重要な一生涯のスキルである。ベネット大学ロースクールの英語教員は、1 年生に対して論理的小論文のトピックを紹介し、5 パラグラフからなる議論の基本構造を教え、内容を練る。この教授法ではiLearn　LMSとClarity　Englishプログラム(英語教育のために作られたソフトウェア)というオンラインプラットフォームを利用し、学生を論文作成チュートリアルに参加させている。本研究は学生がどのように強い主張の構築を理解し、自分のものとするのかについて分析する。研究は実験群と比較群の事前テストと事後テストの結果の比較とトランスクリプトの内容分析によって行われた。被験者はBA　LLB(Hons)の学生120人で実験群と比較群に分けられた。また、本研究は、教室での指導とiLearnフォーラムでのオンライングループディスカッション、そしてClarity Englishが提供するStudy Skill SuccessやPractical Writingによる自習の相対的利点を明らかにしようと試みる。これは学生調査の分析によって行われる。データは事前テストと事後テストにおける小論文のト

ランスクリプトとiLearnでのオンラインディスカッションのトランスクリプト、そして学生調査を含む。本研究は法科の学生に論理的小論文を教えるために利用可能なリソースを用いた最良の指導の実践を評価、デザインする試みである。

Keywords: **Arguments, online tutorials, law undergraduates, LMS, Clarity English**

Lawyers need to have good persuasive and argumentative skills – more than anybody else as it is a life skill for them that they need to acquire with perfection (Viator, 2011; Barnwell, 2015). An aspiring lawyer needs to place a lot of emphasis on learning the ways of presenting cogent arguments in a clear and persuasive manner. Learning about arguments in both spoken and written form provides a strong basis for them in the future to connect it with the cases they would handle in the future. So, what needs to be taught to the students of law first is, *Raise your voice but improve your argument first*. Argument writing is a key skill for law students to master. Almost all components of academic writing involve the construction and justification of an argument for them. While devising the syllabus for an academic writing course for law students, one of the components that cannot be dispensed with is writing a balanced argument in order to get effective persuasion (Bruce, 2002; Ashley et al., 2004; Carr, 2003; Pinkwart et al., 2006). Argumentative writing is a highly important yet lesser researched area for undergraduate education (Pessoa et al., 2017), especially in law courses. Looking at the importance of this component, instructors need to provide maximum practice to students for refining their ability to draft a sound argument. Moreover, it is essential to first develop the understanding among students of the basic structure of an argument and provide essential scaffolds (Øgreid & Hertzberg, 2009) to enable them to appropriate it.

The terms *appropriation* and *scaffolds* have reference to Vygotsky's (1980) socio-cultural theory that states that in order to get students internalise any concept, they need to be provided with some support mechanism or scaffolds so that they can gradually move from their entry level of proficiency to the expected level of outcome and attain appropriation in it. In the present research these support mechanisms that work on interaction, collaboration and self- regulation are the online platforms of i-Learn, i.e., a Moodle-based university learning management system (LMS) (discussion forums in the present research) and online language learning programs offered by Clarity English (CE). CE is a Hong Kong-based group that provides customised programs for online English learning. Bennett University has taken subscription to four of its programs- Study Skills Success (SSS), Practical Writing (PW), Tense Buster, and Road to IELTS. For teaching and learning argumentative skills, both LMS discussion forums and CE programs were used as technological scaffolds providing online writing tutorial sessions. These were utilised to support regular classroom teaching and learning process to assess and design the best teaching practice with the available online resources for teaching argumentative essay writing to law undergraduates.

Literature Review

Technical Scaffolds in Argument Writing

Scaffolding argument/argumentative essay writing with technology has been researched quite a few times in university and school settings (Hoffmann, 2015; Kim, 2018; Li, 2006; Lu & Zhang, 2013; Williams & Beam, 2019; Xu et al., 2019). Latifi et al. (2020) have studied the difference between the students' argumentative writing scores through two different treatments – worked examples and opportunities to collaborate and learn through scripting. The study shows that when opportunities for collaborative writing and peer feedback are provided to students, they learn significantly better than the students who learn through worked examples. Worked examples provide step-by-step details of the writing process to an argumentative essay and then expect the students to write on their own, whereas collaboration and scripting provide issues/topics to the students and allow them to explore and solve those issues in collaboration and then write on their own. The present study also employs two experimental groups given two different treatments – written group discussion (GD) on online forums over the LMS and a customised exercises group. The CE programs focus more on testing the basic understanding of students. The study develops on the future research direction given by Latifi et al. (2020) that use of both qualitative and quantitative approaches will give a better understanding of the difference in the performance level, student responses to the interventions and how students behave, interact during the treatments when they are provided with opportunities for independent learning, collaborate and give feedback.

Methodology

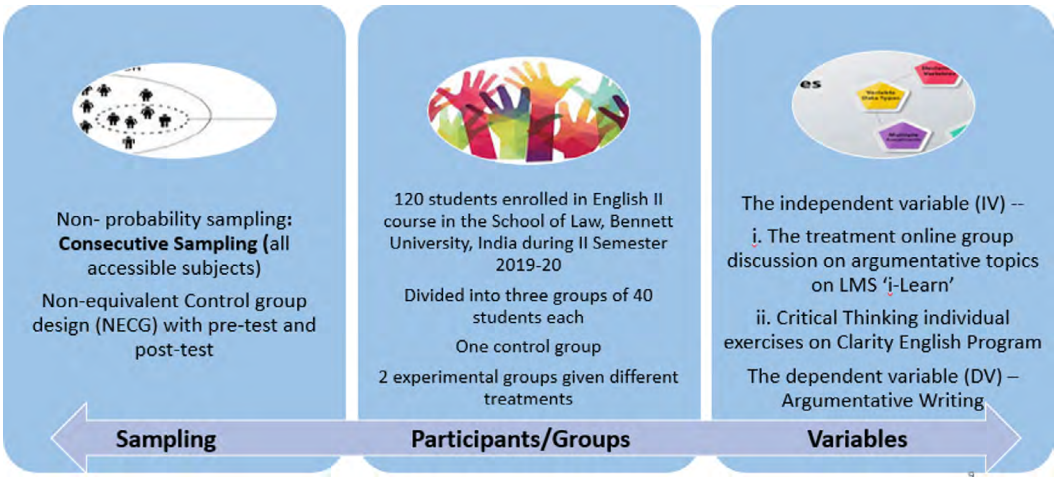
Design

The present study follows a quasi-experimental research design with consecutive sampling. As the design is implemented in educational settings, all available participants were considered (Fife-Schaw, 2006, p. 93; Park & Han, 2018). There were a total 120 students out of which 40 are in the control group and 40 each in the two experimental groups (see Figure 1). All these students were enrolled in the course English II offered as a compulsory course to law undergraduates during the II semester (2019–20) in I year at Bennett University. This course takes up the components of academic writing for law students.

The study focuses on analysing the effect of online tutorials using online LMS forums and CE programs. To see the effect of both types of online tutorials, multiple qualitative and quantitative methods – pre-test / post-test, content analysis and student surveys – have been employed.

Figure 1

Research Design, Groups, and Variables



Before starting to teach the component of argument writing, an entry level proficiency check of students' familiarity with an argument was also done for these students. It was observed that students entering Bennett Law School had negligible exposure to argumentative essays or argumentation until senior secondary level. 80% of the students did not learn that at school. A few students who were active public speakers mentioned frequently debating but oral form of argumentation is little different from written and more structured way of arguing, something that they were not aware of. To confirm it further, simple day-to-day argumentative prompts were given to students in the initial class and they were asked to write for/against the topic and then justify with reasons and evidence. The exercise showed that students were not able to differentiate between reasons and claims, and further that they could not find relevant evidence to support their claims which could directly relate to the reasons that they give to support their claims.

Blended Learning and Scaffolds for Online Tutorials

While teaching basics of argument structure under argumentative essay writing, the students were divided into three groups to test the efficacy of blended learning (Jin et al., 2020) – one control group and two experimental groups. The control group was not given any additional practice apart from classroom discussions on argument structure. Among experimental groups, experimental group 1 was given the treatment of online tutorial GD on debatable topics (see Figure 2) and the activity was named as online written GD and clear instructions regarding the parameters of evaluation were given to them. The students were required to post and centre their discussion on thesis statements, claims, reasons, and evidence, along with basic counterarguments and rebuttals. This was done to make them do focused discussion.

Figure 2

Online GD Topics as Posted on LMS

Some content on internet should be censored.	Suman Luhach	Group A
India is deconstructing its image of a soft nation	Suman Luhach	Group B
Religion is a mitigating factor in India's development	Suman Luhach	Group C

Experimental group 2 was given the task of completing all exercises in the unit of Critical Thinking of SSS and Essays: For and Against of PW among CE programs (see Figure 3).

Figure 3

Clarity English Programs

The exercises in these programs contain mostly objective type application-based questions that focus on testing – understanding of arguments, evaluating evidence, flawed logic, format, brainstorming, stages of writing, writing style and process, and language associated with argument and essay writing.

Both experimental groups participated in online tutorials related to structure and understanding of arguments. The basic difference between the treatments given to both experimental groups lies in two aspects:

1. Nature of the tasks: for group 1, it was an open group discussion with instructional scaffolding and increased scope for peer feedback while for group 2 it was

individual practice sessions on the standardised and uniform online language learning programs.

2. Type of participation: for group 1 participation was instructionally structured, i.e., students had to follow the teacher's instructions regarding the type of posts with flexibility of quantity of content while for experimental group 2 it was highly structured having pre-determined program instructions and questions with no flexibility.

Research Questions

The present study attempts to create two different types of online tutorial environments for practicing and revising argument writing as a supplementary to the physical classroom. The study has five research questions:

RQ1. Is there any significant difference in the mean gain scores of the control group and experimental group 1?

RQ2. Is there any significant difference in the mean gain scores of the control group and experimental group 2?

RQ3. Is there any significant difference in the mean gain scores of experimental groups 1 and 2?

RQ4. What is the nature of online group discussion on argument structure?

RQ5. What is the students' perception about online tutorials?

Data Collection as per the Pedagogical Sequence

In a two-month long teaching and tutorial process of argument structure and essays, students were first collectively taught the basics of argument writing and basic drafting of an argumentative essay for 4 weeks with two lectures per week. At the end of week four, a pre-test of all three groups was conducted (see Figure 4).

During week 5–7 the experimental groups were exposed to two different online tutorial environments – online GD on argument drafting over forums on LMS and CE programs' units on Critical Thinking and Essays: For and Against. A total of five groups of 8 students each (a total of 40) were created for online group discussion for the experimental group 1. Online transcripts of GD and performance sheets of CE programs were saved as records to analyse the GD and writing process as well as students' progression in structured programs of CE. Later, in week 8, a post-test of all three groups was conducted.

Figure 4

The Pedagogical Sequence

Data Analysis

Pre-test and Post-test

Pre-test and post-test analysis were done using an analytical rubric having 9 parameters: attention grabber, thesis statement, claims, evidence, reasoning, coherence, mechanics, conclusion, and documentation. The rubric defined a range of levels from 0 to 3 – 0 (unacceptable), 1 (developing), 2 (accomplished), and 3 (exemplary) – for befitting completion of each component of an argumentative essay. Total number of items in the rubric is nine that sets the highest score for the rubric at 27. For checking the validity and reliability of the rubric, it was rated by three academic writing experts. The experts agreed on 76.33% level of ratings. Cohen's Kappa Coefficient of reliability was .65 (p <.0.001) which is considered a good level of inter-rater reliability (Bakeman & Gottman, 1997).

Variables for pre-test and post-test analysis:

- The independent variable (IV): Online GD (Group 1); Exercises on CE Programs (Group 2)
- The dependent variable (DV): Argumentative essay writing

Student Surveys

Student surveys was taken separately for both experimental groups. The surveys for both group 1 and group 2 consisted of five open-ended questions. The questions sought students' opinions on their liking, disliking, and suggestions of the two online tutorial methods.

Content Analysis of Online GD Transcripts

Online GD transcripts provided the data to analyse the nature of written discussion on argument structure. It was done through content analysis of student posts on online discussion forums. The content analysis consists of defining a concept and tallying its presence (Busch et al., 2012). The defined concepts included components of an argument – thesis, claims, reasons, evidence, counterclaims, and rebuttals. Kappa statistic for inter-rater reliability was 0.69 ($p < 0.001$) which is a good level of inter-rater reliability.

Results and Discussion

Pre-test and Post-test

The scores of pre-test and post-test were analysed by descriptive and inferential statistics. The descriptive statistics analysed the tests on the mean, standard deviation, and percentages. The mean scores of the control group, experimental group 1, and experimental group 2 were assessed to see if any statistical difference exists in the performance of argumentative writing of students.

Descriptive Statistics

The results show that students in the experimental group 1 and 2 performed much better in the post-test than in the pre-test for the composite scores of argumentative essay writing than the control group. In the composite scores of the pre-test and post-test of the experimental group 1 (see Table 1), the mean score for pre-test was 16.33 and 23.33 for the post-test. The percentage increase in the mean score for experimental group 1 is 42.86%. In the composite scores of pre-test and post-test of experimental group 2 (see Table 1), the mean score for pre-test was 15.73 and 19.35 for the post-test.

Table 1

Descriptive Statistics for Experimental and Control Group

Group		Mean	Std. Deviation	% change
Control Group	Pre-test	15.43	4.358	10.30% increase in post-test score
	Post-test	17.02	4.854	
Experimental Group 1	Pre-test	16.33	4.221	42.86% increase in post-test score
	Post-test	23.33	3.89	
Experimental Group 2	Pre-test	15.73	3.464	23% increase in post-test score
	Post-test	19.35	4.035	

The percentage increase in the mean score for experimental group 2 is 23%. For the control group, the mean score in pre-test was 15.43 and 17.02 in the post-test. The percentage

increase in the mean score was 10.30% which is much smaller than the percentage change of 42.86% and 23% for the experimental group 1 and 2, respectively.

Inferential Statistics

The research hypothesis 1 for analysing pre-test and post results is: The average score of argumentative essay writing for students who receive online GD treatment is greater than the average score of students who did not receive the treatment. The results from an independent samples *t*-test indicate that students who received recreated online GD treatment (see Table 2) (M = 23.33, SD = 3.89, N = 40) scored higher than students who did not receive the treatment (M = 17.02, SD = 4.854, N = 40). Cohen's effect size value (d = 1.434) suggested a 'large' effect size and high practical significance.

Table 2

Group Statistics I

Group	*N*	Mean	SD	*df*	*T*	*p*-value	Cohen's *d*	Remark
Experimental Group 1	40	23.33	3.89	78	–6.489	0.000	1.434	significant
Control Group	40	17.02	4.854					

The research hypothesis 2 for analysing pre-test and post results is: The average score of argumentative essay writing for students who received online CE tutorial treatment is greater than the average score of students who did not receive the treatment. The results from an independent samples *t*-test indicate that students who received recreated online DC treatment (see Table 3) (M = 19.35, SD = 4.035, N =40) scored higher than students who did not receive the treatment (M = 17.02, SD = 4.854, N =40). Cohen's effect size value (d = 0.521) suggested a 'medium' effect size and practical significance.

Table 3

Group Statistics II

Group	*N*	Mean	SD	*df*	*T*	*p*-value	Cohen's *d*	Remark
Experimental Group 2	40	19.35	4.035	78	–2.329	0.023	0.521	significant
Control Group	40	17.02	4.854					

The research hypothesis 3 for analysing pre-test and post results is: The average score of argumentative essay writing for students who receive online GD treatment is greater than the average score of students who received online CE tutorial treatment. The results from an independent samples *t*-test indicate that students who received recreated online DC treatment (see Table 4) (M = 23.33, SD = 3.89, N =40) scored higher than students who received online CE tutorial treatment. (M = 19.35, SD = 4.035, N =40). Cohen's effect size value (d = 1.004) suggested a large effect size and high practical significance.

Table 4

Group Statistics III

Group	*N*	Mean	SD	*df*	*T*	*p*-value	Cohen's *d*	Remark
Experimental Group 2	40	23.33	3.89	78	−4.491	0.000	1.004	significant
Control Group	40	19.35	4.035					

Pre-test and post-test results indicate that both experimental groups 1 and 2 performed significantly better in post-test as compared to the control group. Between experimental groups, the performance of students of experimental group 1 is significantly better than the students of experimental group 2. This implies that though the students of both experimental groups gained better understanding and appropriation in argument writing, students engaged in online GD benefitted more than students who were engaged in highly structured CE programs of critical thinking and essays for and against. This also implies that when students are given flexibility and freedom of expressing themselves, while remaining careful about the integral elements and structure of an argument, they are more likely to self- regulate their learning process and subsequently improve.

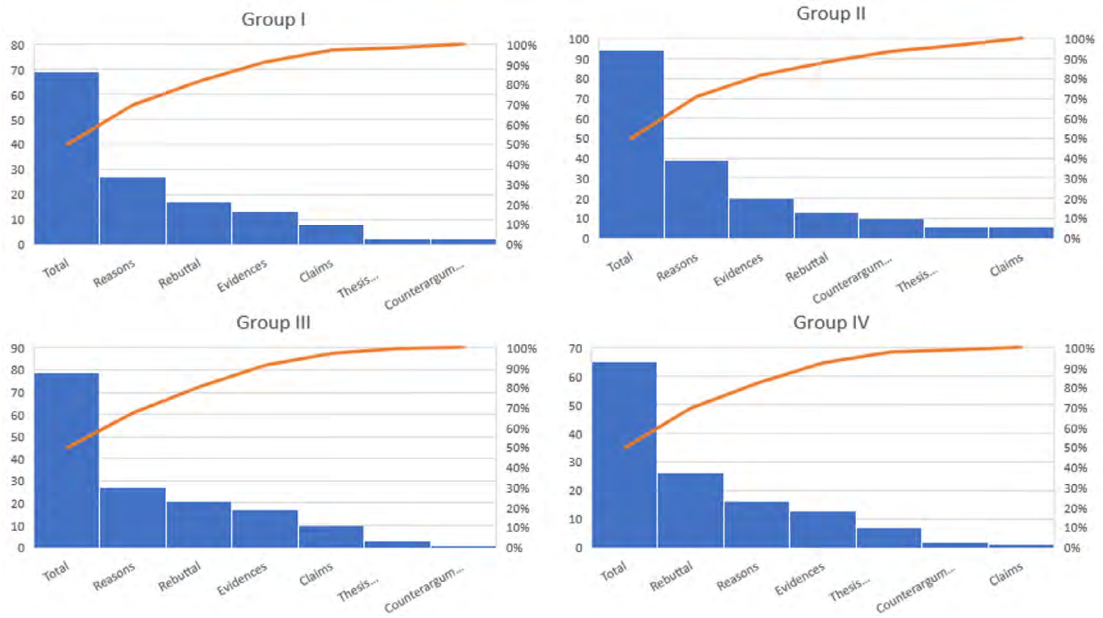
Nature of Online Group Discussion on Argument Structure

Experimental group 2 was involved in structured objective type exercises on CE programs, hence the nature of their learning process of argument writing could be understood only through a student survey whereas experimental group 1 did online written GD for which the transcripts were saved by the instructor to analyze the nature of discussion on argument structure. This was done by doing content analysis of the transcripts of the four groups made for GD on online forums on Bennett University's i-Learn LMS. This analysis of the nature of GD was done to see the elements most discussed, their strength and direction.

Content analysis was done by showing the frequency distribution of posts through Pareto charts. Pareto charts are made mainly to show that 80 percent of the cause of something is just 20 percent of the factors. Here, the charts show how maximum number of posts were contributed by just three components of an argument which is similar across the four groups made for GD on different topics (see Figure 5).

Figure 5

Argument Component Frequency Distribution



The graphs in Figure 5 clearly reflect that the maximum number of posts contributed for reasons, rebuttal, and evidence for an argument once the thesis statement and claim are defined and decided among group members during the GD. The pattern was similar across groups. Surprisingly, even the number of counterarguments was smaller, which implies that students focused more on the justification of their stand through multiple reasons and evidence.

Student Surveys

Survey Results for Online GD on LMS i-Learn

Student surveys for both experimental groups obtained mixed reactions. For experimental group 1, students acknowledged that online GD helped them focusing, refining, and justifying. Peer feedback was also taken as a crucial aspect while drafting and refining the argument: "It has helped me in learning how to frame an answer and present it so that it remains relevant and adds on to the discussion."

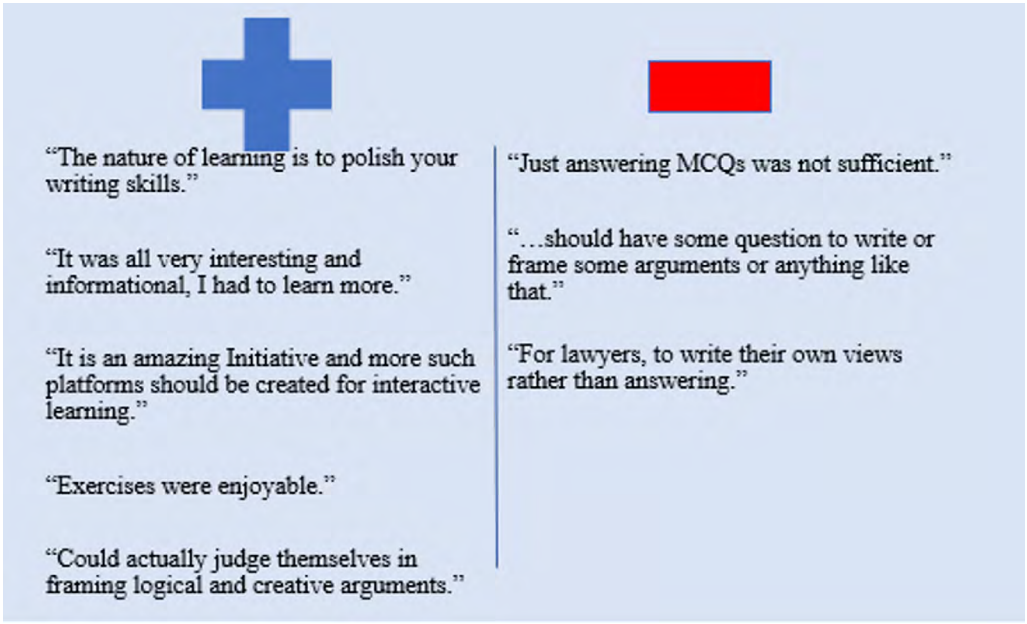
The students considered rebuttals as very effective in refining their argument. The only concern that the students expressed while discussing was that sometimes students got into heated debates while trying to disprove each other by blaming each other, and they deviated from the actual task. Some students expressed the desire of increased involvement of the instructor in the discussions so that such distractions from the actual task of argument discussion would not happen and students would start giving timely responses: "Some people took more time to respond to the arguments that were against my claims, so it was difficult to have a debate/a discussion actively."

Survey Results for Online Tutorial Based on Clarity English Programs

The CE program based online tutorial also entailed mixed reaction from the students. They added that the tasks were structured in such a manner that they helped students in polishing their writing skills. A few students found those informational and interesting to do and suggested to incorporate more such exercises. On the other hand, some students felt that the objective-type nature of questions was not sufficient to hone writing skills as such questions do not provide any space for free expression of ideas to students. They suggested some detailed writing practise to be given in place of very short answer questions like MCQs, fill the blanks, or one-word answers. They agreed to have some enhanced understanding of the flawed logic and sound argument but that did not allow them to write and correct their own arguments (see Figure 6).

Figure 6

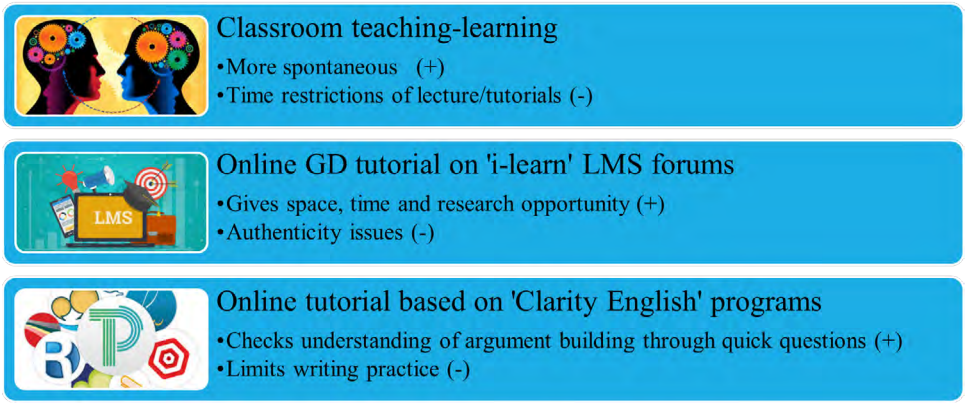
Student Response on Clarity English Programs for Online Tutorial



The results of pre-test post-test, content analysis, and student surveys collectively have thrown light on the differences, benefits, and drawbacks of the three different ways of engaging students in teaching-learning process of argument writing in argumentative essays. The study tried to do a comparative analysis of classroom teaching in physical settings, online group discussions on i-Learn LMS forums and individual practice sessions involving only one student at a time on SSS and PW program offered by CE. It was found that in the physical classroom students tend to be more spontaneous and come with very interesting and original ideas but those ideas could be only strengthened through further research and alleviation of time and space constraints of a physical classroom (see Figure 7).

Figure 7

Comparative Advantages/Disadvantages of Different Ways of Argument Writing Teaching-learning Process



The online tutorial of GD on argumentative topics gives the convenience of time and space for doing adequate research on the given topics. But there is a possibility of having some authenticity issues about the content posted by students. On the other hand, online tutorials based on CE programs check the understanding of the students through some quick questions but have a drawback of limiting the writing practice, feedback, and internalising the writing process of arguments.

Conclusion

The present study attempts to analyse the development in the understanding of the students in framing a strong argument and movement towards attaining appropriation in it. Both experimental groups showed some improvement in their understanding of argument structure and they performed better in their post-test results. But out of the two groups, experimental group 1 showed better performance. It implies that complimentary and spontaneous contribution by the students engaged in online GD worked better on their understanding and appropriation than the structured exercises of CE programs. This happened as the students, in online GDs, were subjected to an environment where they had to become critics of not only their own learning but also of their peers' learning. The study also aimed to devise the best pedagogical practice for honing argumentative writing skills among law students. While analysing the comparative advantages and disadvantages of teaching learning process during physical classroom teaching, online writing tutorials on LMS discussion forums, and CE programs, it was observed that all three ways have their own benefits. In agreement with Latifi et al. (2020), the present study also suggests that both approaches have their own benefits, and it is necessary to analyse how those could be used in honing different aspects of students' writing. These different methods, if blended with proper planning, can help students better appropriate the writing of arguments. Spontaneous quick discussions can be fostered during classroom teaching which could be taken further to an online platform for GD for getting insight into argument development with proper structure. This would also address the problem of time crunch during regular lectures. Clarity English

programs comprise some standardised exercises which could help students to quickly estimate their understanding of critical thinking, flaws in logic, argument structure, and essay structure. These exercises can be incorporated into blended learning as tools for checking students' in depth understanding in the form of quizzes. Certain issues found in online tutorials like the issue of authenticity of posts by the students with respect to the originality of idea and proxy participation during online GDs in asynchronous mode could be a subject of further research as a valuable addition to the pedagogy.

References

Ashley, K.D., Desai, R., & Levine, J.M. (2002). Teaching case-based argumentation concepts using dialectic arguments vs. didactic explanations. In *International conference on intelligent tutoring systems* (pp. 585–595). Springer. **http://doi.org/10.1007/3-540-47987-2_60**

Bakeman, R., & Gottman, J.M. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge University Press.

Barnwell, D. (2015). *Rhetoric and Law: How Do Lawyers Persuade Judges? How Do Lawyers Deal with Bias in Judges and Judging?* [Master's thesis, Georgia State University]. **https://scholarworks.gsu.edu/english_hontheses/10**

Bruce, N. (2002). Dovetailing language and content: Teaching balanced argument in legal problem answer writing. *English for Specific Purposes*, *21*(4), 321–345. **http://doi.org/10.1016/S0889-4906(01)00040-0**

Busch, C., De Maret, P.S., Flynn, T., Kellum, R., Le, S., Meyers, B., Saunders, M., White, R., & Palmquist, M. (2012). *Content Analysis*. Writing@ CSU. Colorado State University. **https://writing.colostate.edu/guides/guide.cfm?guideid=61**

Carr, C.S. (2003). Using computer supported argument visualization to teach legal argumentation. In P.A. Kirschner, Shum S.J. Buckingham, & C.S. Carr (Eds.), *Visualizing argumentation* (pp. 75–96). Springer. **http://doi.org/10.1007/978-1-4471-0037-9_4**

Fidel, G., & Cantoni, L. (2004). *Think like a lawyer: The art of argument for law students*. Xlibris Corporation.

Fife-Schaw, C. (2006). Quasi-experimental designs. *Research methods in psychology*, pp. 88–103. Sage.

Hoffmann, M.H. (2015). Changing philosophy through technology: Complexity and computer-supported collaborative argument mapping. *Philosophy & Technology*, *28*(2), 167–188. **http://doi.org/10.1007/s13347-013-0143-6**

Jin, T., Su, Y., & Lei, J. (2020). Exploring the blended learning design for argumentative writing. *Language Learning & Technology, 24*(2), 23–34. **http://hdl.handle.net/10125/44720**

Kim, N.Y. (2018). Corpus technology and video technology: Effects on EFL argumentative writing. *The New Korean Journal of English Language and Literature, 60*(2), 195–217

Øgreid, A.K., & Hertzberg, F. (2009). Argumentation in and across disciplines: Two Norwegian cases. *Argumentation*, *23*(4), 451. **http://doi.org/10.1007/s10503-009-9162-y**

Latifi, S., Noroozi, O., & Talaee, E. (2020). Worked example or scripting? Fostering students' online argumentative peer feedback, essay writing and learning. *Interactive Learning Environments*, 1–15. **http://doi.org/10.1080/10494820.2020.1799032**

Li, J. (2006). The mediation of technology in ESL writing and its implications for writing assessment. *Assessing Writing*, *11*(1), 5–21. **http://doi.org/10.1016/j.asw.2005.09.001**

Lu, J., & Zhang, Z. (2013). Scaffolding argumentation in intact class: Integrating technology and pedagogy. *Computers & Education*, *69*, 189–198. **http://doi.org/10.1016/j.compedu.2013.07.021**

Øgreid, A.K., & Hertzberg, F. (2009). Argumentation in and across disciplines: Two Norwegian cases. *Argumentation*, *23*(4), 451. **doi: 10.1007/s10503-009-9162-y**

Park, Y., & Han, K. (2018). Development and evaluation of a communication enhancement program for people with chronic schizophrenia: A quasi-experimental pretest-posttest design study. *Applied Nursing Research*, *42*, 1–8. **http://doi.org/10.1016/j.apnr.2018.04.007**

Pessoa, S., Mitchell, T.D., & Miller, R.T. (2017). Emergent arguments: A functional approach to analyzing student challenges with the argument genre. *Journal of Second Language Writing*, *38*, 42–55. **http://doi.org/10.1016/j.jslw.2017.10.013**

Pinkwart, N., Aleven, V., Ashley, K., & Lynch, C. (2006). Toward legal argument instruction with graph grammars and collaborative filtering techniques. In *International Conference on Intelligent Tutoring Systems* (pp. 227–236). Springer. **http://doi.org/10.1007/11774303_23**

Viator, J.E. (2011). Legal education's perfect storm: Law students' poor writing and legal analysis skills collide with dismal employment prospects, creating the urgent need to reconfigure the first-year curriculum. *Cath. UL Rev.*, *61*, 735.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Williams, C., & Beam, S. (2019). Technology and writing: Review of research. *Computers & Education*, *128*, 227–242. **http://doi.org/10.1016/j.compedu.2018.09.024**

Xu, Z., Banerjee, M., Ramirez, G., Zhu, G., & Wijekumar, K. (2019). The effectiveness of educational technology applications on adult English language learners' writing quality: A meta-analysis. *Computer Assisted Language Learning*, *32*(1–2), 132–162. **http://doi.org/10.1080/09588221.2018.1501069**

Appendix A

Rubric for Pre-test and Post-test

S. No.		Marks 0 (unacceptable)	1 (developing)	2 (accomplished)	3 (exemplary)	Total
1	**Attention grabber**	Not found/ irrelevant	Attention grabber needs refining	Attention grabber attracts attention and connects to the purpose	Impressive and perfectly connecting attention grabber	
2	**Thesis statement**	No thesis statement/ irrelevant statement	The thesis statement needs refining	Clear thesis statement	Very well drafted, clear thesis statement	
3	**Claims**					
4	**Evidence**	No evidence/ irrelevant evidence	Evidence found but not consistently with all claims	Evidence found with all claims	Very strong evidence	
5	**Reasoning**	No supporting reasons with claims/ irrelevant reasons	Reasons found but not consistently with all claims	Reasons found with all claims	Very strong reasons with all claims	
6	**Coherence**	No thematic connection	Maintained at some places	Shows thematic unity	Very systematic and well organised	
7	**Mechanics**	No attention paid	Considered at some points	Rare error	No error	
8	**Conclusion**	Ineffective or no conclusion.	Not emphasized main points and lacks clarity.	Emphasized main points with and clarity.	Emphasized main points with effective clinchers and summarized with clarity	
9	**Documentation**	Not found	Not in proper format/ irregular	Done for all references	Done for all in accurate format	
10	**TOTAL**					

Appendix B

Student Survey for Online Group Discussion on LMS i-Learn

Dear Student, you are requested to share your experience of online GD and argument writing on online forums. Your honest responses will help in determining the efficacy of the assignment. Please respond to all sections of the questionnaire. All the information will be kept confidential.

Qs. 1. Share your experience of online group discussion.
Qs. 2. Has peer feedback helped you in refining your argument or any component of argument – claim, reason, evidence? If yes, state how?
Qs. 3. What problem did you face in online written discussion?
Qs. 4. Any suggestion for improvement.
Qs. 5. Anything else you want to share.

Appendix C

Student Survey for Online Tutorial Based on Clarity English Programs

Dear Student, you are requested to share your learning experience of *Critical thinking* of SSS and *Essays: for and against* of PW among CE programs. Your honest responses will help in determining the efficacy of the assignment. Please respond to all sections of the questionnaire. All the information will be kept confidential.

Qs. 1. Share your experience of online tutorial based on Clarity English Programs.
Qs. 2. Have structured exercises helped you in refining your argument or any component of argument- claim, reason, evidence? If yes, state how?
Qs. 3. What problem did you face in online tutorial based on Clarity English Programs?
Qs. 4. Any suggestion for improvement.
Qs. 5. Anything else you want to share.

Author's Bio

Suman Luhach is an Assistant Professor of English in School of Law, Bennett University, India. She has a PhD in English from BITS Pilani, India. Her research and interest areas include academic and advanced writing skills of college students, online collaborative language learning, CALL, ELT, legal English, and literary theories.

https://doi.org/10.37546/JALTSIG.CALL2020.8

8

A Comparison of the Affective Affordances of a Static and Interactive VR System on Learner FLA and Motivation

Koichi Shibata and James York, Tokyo Denki University

Abstract

This paper introduces a virtual reality (VR) system which was designed to promote English speaking proficiency as learners carry out collaborative information gap tasks. In a former study, a simpler system was developed to explore the effect of modality on learners' foreign language anxiety (FLA) where results suggested that anxiety was statistically significantly lower in the VR environment compared to a voice and video chat system. However, of three key affordances – presence, interactivity, and autonomy – the previous system only focused on presence. The current system features an interactive component also. In this paper, we present results of a study which compared the two systems (presence-only versus interactive system) with the aim of answering the question: Does more-fully utilizing the affordances of VR lower or increase learners' FLA?

In a counterbalanced design, 30 participants (15 pairs) completed a spot-the-difference task in two different VR environments: static-VR (former system) and interactive-VR (current system). Results of a post-experimental questionnaire suggested that there was no difference in participants' FLA for the two domains. However, a significant difference was found in terms of ease of communication and enjoyment which favored the interactive-VR mode. Additionally, compared to predictions that the interactive-VR task would be more cognitively demanding, it was considered simpler than the static-VR task. This suggests that using more of the affordances of VR by increasing interactivity further may make the embodied experience more life-like and therefore increase opportunities for learning. This paper introduces the system, implications for researchers and teachers, and future research directions.

本論文では、共同でinformation gap taskを実行する際に、英会話能力を促進するように設計された仮想現実（VR）システムを紹介します。以前の研究では、コミュニケーション方法が学習者の外国語不安感（FLA）に及ぼす影響を調査するために、シンプルなシステムが開発され、学習者の不安は、音声およびビデオチャットシステムと比較して、VR環境では統計的に有意に低いことが示唆されました。ただし、3つの主要な要素（没入感、対話性、および自律性）のうち、以前のシステムは没入感のみに焦点を合わせ

ていました。現在のシステムは、インタラクティブな機能を追加しました。この論文では、VRの要素をより十分に活用することで学生のFLAが低下または増加するのかという調査を行いました。

Keywords **: SCMC, VR, interactivity, motivation, anxiety**

Virtual reality (VR) exists along a continuum of augmented, mixed, and virtual realities, which refers to the level of immersion into a virtual environment each system provides (Hawkinson et al., 2017). Although originally conceptualized in the 1960s (see Sherman & Craig, 2018), VR did not become a popular consumer product until the last decade with the gradual and incremental improvements made in technology (both hardware and software). Stein (2019) provides a succinct history of developments over the last 10 years which include:

- Head-mounted display systems which require a desktop computer (HTC Vive, Oculus Rift)
- Augmented reality glasses (Google Glass)
- Augmented reality games which utilize smartphones GPS and gyroscope functions (Pokémon Go being one of the most popular games)
- Smartphone VR systems (Google Cardboard)
- Stand-alone VR systems (Oculus Quest)

In sum then, the development of VR technologies is growing, and becoming more affordable for consumers who are being provided with systems that are both cheaper and easier to use. Along with this trend of VR entering the consumer market then, research within CALL is also beginning to utilize the affordances of this new technology for language learning purposes.

The research outlined in this paper has the aim of developing a system which can be used to connect Japanese learners with native English speakers in a virtual environment to complete focused, closed-goal tasks (Ellis, 2003). The system has a specific focus on oral communication as this skill is most neglected or difficult to practice in classroom environments void of native speakers. The current paper introduces results of a study which sought to understand how the cognitive demands of completing a collaborative information gap task in a VR system could affect learners' foreign language anxiety (FLA) and motivation towards studying a foreign language. As this is a preliminary test of the system, native speakers were not a part of the study.

The reason that we focus on FLA is that it is known to negatively affect learner motivation to study a second or foreign language (Hewitt & Stephenson, 2012). It is considered one of the most hindering barriers to successful L2 acquisition, affecting a learner's motivation, and self-efficacy in the L2 (Dewaele & Thirtle 2009; Horwitz et al., 1986). The reason we focus on an oral collaborative task is because the context of this study is Japan, where the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) promote the development of practical English skills including being able to express oneself in English (MEXT, 2014, p.3). However, as Japan is a homogeneous nation (Ebuchi & Yokota, 2019) there are few opportunities for residents to interact with native speakers in face-to-face communication. Certainly, the Internet and other forms of computer-mediated

communication (CMC) provide opportunities for learners to communicate with native speakers, but VR holds the most promise in terms of allowing for an embodied experience which has been called a "hybrid mode of communication which may offer the benefits of both face-to-face and computer-mediated modalities" (York et al., 2020, p.2).

Literature Review

Computer-mediated Communication in Language Learning Contexts

Before outlining the research on VR in language learning contexts, this section introduces the wider topic of computer-mediated communication (CMC) studies within CALL. There is a growing body of literature exploring the cognitive and affective affordances of CMC, which is further divided into asynchronous (ACMC) and synchronous (SCMC) modes. Here, the focus is on SCMC.

In relation to the present study, there are several meta-analyses on the effect of SCMC in language learning contexts (Lin, 2014; Aslan & Ciftci, 2019). Ziegler's (2016) meta-analysis examined 14 studies which utilized SCMC as a mode for interaction between learners. Findings of the meta-analysis suggested that both face-to-face (FTF) and SCMC interactions produced significant, positive effects on learners' language development. This confirms the notion that interaction – and in turn the Output Hypothesis – is beneficial to language development, but also that SCMC is not inferior to FTF communication. Furthermore, there was a small, positive advantage for SCMC over FTF for written production. However, this may be a moot point when considering the types of SCMC modes analyzed in this meta-analysis. SCMC does not imply oral interaction, and indeed, only three of the 14 included studies incorporated oral SCMC. Thus, if the SCMC mode is written, one would expect there to be an improvement in written output in comparison to FTF communication, which by necessity is oral.

Regarding the affective affordances of SCMC, text-based SCMC has been shown to reduce FLA in comparison to FTF communication (Abrams, 2003; Tudini, 2007). Furthermore, studies have shown that SCMC can promote more equal turn-taking opportunities (Warschauer, 1996) and improved willingness to communicate among peers (Reinders & Wattana, 2014; Yanguas & Flores, 2014) as well as with native speakers (Jauregi et al., 2012; Iino & Yabuta, 2015). Melchor-Couto (2017) conducted a study comparing the FLA of an experimental group who completed oral tasks in a virtual world to a control group that conducted the same tasks face-to-face. Results suggested that the FLA of participants in the experimental group went down over multiple meetings within the virtual world. This was attributed in part to the anonymity afforded by avatar-based communication (see also Dickey, 2005).

The Use of VR in Language Learning

As mentioned in the introduction to this paper, VR is currently a hot topic in CALL and broader educational fields (Liu et al., 2017). Several papers have explored hypothetical

applications of VR and related technologies in language learning contexts (Bonner & Reinders, 2018; Hawkinson, et al., 2017; Alizadeh, 2019). However, there are few empirical studies on the cognitive or affective affordances of the technology. Of those that exist, there is an emphasis on exploring VR's effect on vocabulary acquisition (Cheng et al., 2017; Legault et al., 2019) where results claim successful vocabulary learning was due to the immersive nature of the environment, and provision of culturally relevant interactions.

York et al. (2020) compared the effect of three modes of SCMC on learners' FLA: oral SCMC, video SCMC and VR. All three modes were effective in reducing learners' FLA, but compared to oral and video SCMC modes, participants considered VR the most natural environment to communicate in, as well as being the most fun, and the most effective environment for language learning. However, limitations of the study were that the affordances of the VR domain were specifically *not* utilized to unify the complexity of the three chosen modes. The current study may be considered an extension, where an interactive element is added to the previous system. However, the efficacy of this new system in reducing learners FLA is unknown, hence the impetus for this study.

Task Complexity and Cognitive Demand

Robinson's (2011) Triadic Componential Framework for task conditions is a robust framework for assessing and making predictions regarding the *cognitive complexity* of language learning tasks, also known as the *Cognition Hypothesis*. It was developed over a number of iterations (see Robinson, 2000; 2005) and validated in several studies (Jackson & Suethanapornkul, 2013; Sasayama, 2016; York, 2019). The framework attempts to provide practitioners with the ability to predict the complexity of a language task and is separated into three major categories: task complexity, conditions and difficulty. For brevity, this paper focuses on the *task complexity* dimension. As can be seen in Table 1, task complexity is bifurcated into resource-directing and resource-dispersing dimensions. The addition of + and – signs also indicate that task complexity can be instrumentalized by a task designer by increasing or decreasing a certain condition. As a concrete example, increasing the number of elements that a learner has to manipulate in a task (*+few elements*) would increase its complexity along the resource-directing dimension. Thus, the complexity of a task is determined by the cognitive demands placed on learners.

Table 1

Robinson's Triadic Componential Framework (2011)

Task complexity		Task conditions		Task difficulty	
Resource-directing	**Resource-dispersing**	**Participation**	**Participant**	**Affective variables**	**Ability variables**
+/– Here-and-Now	+/– Planning	+/– Open solution	+/– Same proficiency	H/L Openness	H/L Working memory
+/– Few elements	+/– Prior knowledge	+/– One-way flow	+/– Same gender	H/L Control of emotion	H/L Reasoning
+/– Spatial reasoning	+/– single task	+/– Convergent solution	+/– Familiarity	H/L Task motivation	H/L Task switching
+/– Causal reasoning	+/– Task structure	+/– Few participants	+/– Shared content knowledge	H/L Processing anxiety	H/L Aptitude
+/– Intentional reasoning	+/ Few steps	+/– Few contributions needed	+/– Equal status and role	H/L Willingness to communicate	H/L Field Independence
+/– Perspective-taking	+/– Independency of steps	+/– Negotiation not needed	+/– Shared cultural knowledge	H/L Self-efficacy	H/L Mind-reading

Robinson argues that cognitive complexity may have a positive effect on learners' output accuracy and complexity, but a negative effect on fluency as learners focus on producing complex language. Subsequently, due to the extra focus required, acquisition of the target language may also be improved. In relation to the current study, affectively, the increase in cognitive task complexity may have a negative effect on learners' motivation as they become overloaded by the cognitive demands of a task (Sweller, 2005; Paas et al., 2005; van Gog & Paas, 2008).

The topic of task complexity and cognitive demands has relevance to the present study in that we are testing two VR systems which, based on Robinson's Triadic Componential Framework, may be considered more or less complex environments. Thus, in this paper, we are addressing the question of how incorporating interactive elements into a VR language learning system may affect learner performance due to differing task complexities.

Methodology

The aim of this study is to uncover the benefit or hindrance of an interactive VR system on learners' FLA and motivation towards learning.

Participants

The study was carried out at a private university in Japan. 30 participants (15 pairs) volunteered to take part in the study. Their mean age was 20.9 (*SD* = 1.54). All participants were native speakers of Japanese. 27 of the participants were male and three were female which is not uncommon for this university (Table 2). In order to avoid a familiarity effect,

participants were paired with partners that they meet for the first time when undertaking the VR tasks.

Table 2

Participant Details

Baseline Characteristic		n	%
Age			
	18	1	3.3
	19	4	13.3
	20	8	26.7
	21	7	23.3
	22	3	10.0
	23	6	20.0
	24	1	3.3
Gender			
	Male	27	90
	Female	3	10

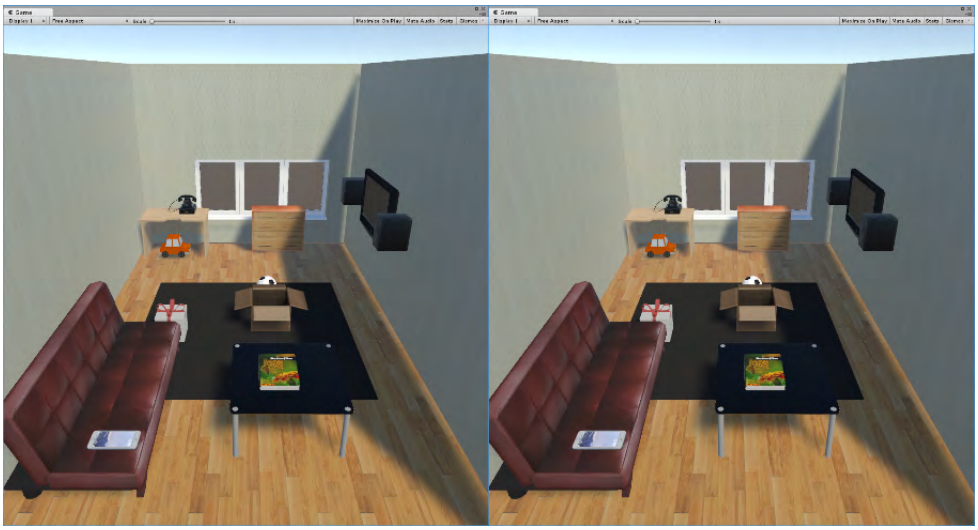
Instruments

VR system

Two VR systems were employed in this study. Both environments were designed as spot-the-difference information gap tasks to be completed by pairs of learners. Both systems present participants with their own doll house which features a number of items typically found in a living room. To create a spot-the-difference activity, three items appear in different places. Figure 1 shows this as the football, the car, and the magazine.

Figure 1

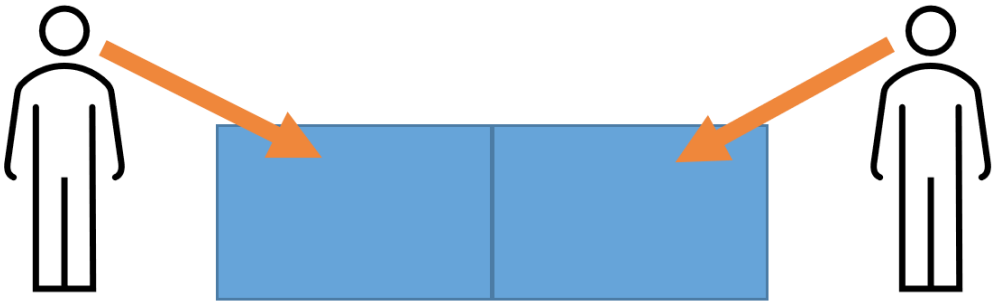
Example of the Doll Houses Presented to Each Participant



Participants are able to see each other standing behind their respective doll house, but not the content, as the doll houses were placed back to back (Figure 2).

Figure 2

Graphical Representation of the VR System Featuring Both Participants



Using the HTC Vive head-mounted display allowed for an immersive VR experience for both systems. Participants were able to look around the environment by moving their head and gesture by moving the controllers in their hands (Figure 3).

Figure 3

A Participant Gesturing with the Controller

The difference between the two systems is that in the previous *static-VR* system participants can only interact with the doll house by looking at it. They cannot move objects in the room. In this way, they must decide when they have found the three differences between their two rooms and end the task. With the current, *interactive-VR* system, participants can pick up and move objects in their room (Figure 4). Additionally, once participants have moved objects into the same position in both doll houses, a message is shown to tell them that they have completed the task.

Figure4

A Screenshot of a Participant Moving an Object in the Interactive-VR system

In summary then, the interactive-VR system was developed as an extension to the static-VR system. A comparison is shown in Table 3.

Table 3

A Comparison of the Static and Interactive VR Systems

Static-VR	Interactive-VR
Doll house	Doll house
Gestures via controllers	Gestures via controllers
	Moveable objects
	Notification when the task is complete

Pre-task

A pre-task was employed to prime learners for the main spot-the-difference task in the VR environments (Appendix A). It featured questions about living room vocabulary, prepositions, and a sample spot-the-difference task. The worksheet was created with images taken from the VR environment. Dyads were allocated 15 minutes to complete the pre-task.

Main Task

The main task was carried out using the VR systems described above. Participants were placed in separate rooms and given instructions on how to use the head-mounted display and controllers. Discord was utilized as the CMC tool, allowing participants to communicate orally as they carried out the task.

Although the two tasks are comparable, the interactive-VR system adds both interactivity (virtual tangibility) and feedback on correct task completion. In terms of Robinson's (2011) Triadic Componential Framework then, the interactive-VR system is considered to be of higher task complexity due to manipulations along the following dimensions:

- **Spatial reasoning:** participants must understand the virtual space to move the objects into the correct places.
- **Prior knowledge:** if participants are not familiar with VR, controllers, and headset, an additional cognitive demand is placed upon them by requiring them to pick up and move objects.
- **Few steps:** There are more steps to complete (moving items into place).

However, as the VR environment is fully immersive and an embodied experience, participants may find the interactive-VR activity intuitive and natural, as found in York, (2020) where a VR SCMC environment was considered an easier mode of communication than both oral and video SCMC.

Reduced FLCAS

A reduced version of Horwitz et al.'s (1986) foreign language classroom anxiety scale (FLCAS) was used in this study to measure participants' FLA at the pre- and post-experiment stages. The reduced version was based on Melchor-Couto (2017; 2018) which features only seven questions:

1. I never feel quite sure of myself when I am speaking in my foreign language class.
2. I do not worry about making mistakes in language class.
3. I start to panic when I must speak without preparation in language class.
4. In language class, I can get so nervous I forget things I know.
5. Even if I am well prepared for language class, I feel anxious about it
6. I feel confident when I speak in foreign language class.
7. I am afraid that the other students will laugh at me when I speak the foreign language.

Post-experiment Questionnaire

Upon task completion, the participants completed the FLCAS and were also asked to give their perceptions of learning within each environment (Table 34). Five measures were utilized based on Satar and Ozdener (2008) and were weighted from 1 to 10, 10 being a strong indication of agreement and 1 disagreement with each statement. In addition, five measures were also included to understand how each environment affected participants' cognitive load, such as in deHaan et al. (2010). Responses to these statements were then used to ascertain the accuracy of task complexity predictions. For example, Statement 1 is an indicator of the cognitive demands of a system where participants' cognitive capacity may be overloaded by the system meaning that they have no capacity left to attend to learning goals (Lim et al., 2006).

Table 4

Measures Used in the Post-Experiment Questionnaire

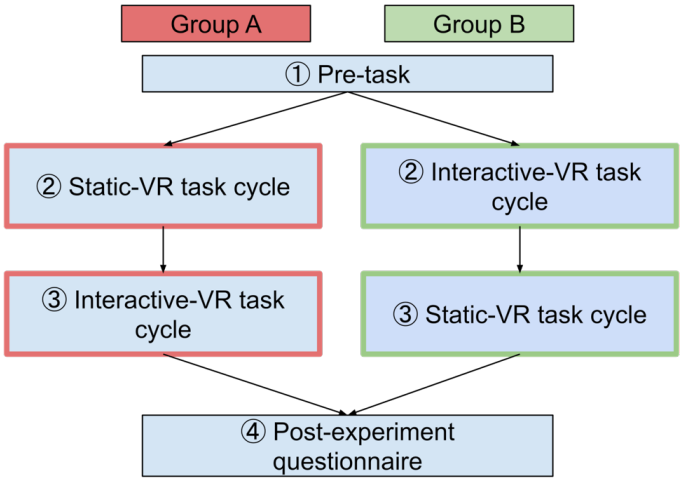
Number	Statement
1	It was easy to speak English when using this system.
2	It was fun to learn with this system.
3	It was easy to complete the task with this system.
4	I think this system is an effective way of learning English.
5	I did not feel anxious learning English with this system.

Procedure

The experiment was a counterbalanced, repeated-measures design where participants completed a spot-the-difference task using both systems. The procedure can be seen graphically in Figure 5. The 15 pairs were assigned to either Group A or Group B. First they both completed the pre-task (1). As part of the pre-task, participants also answered the FLCAS questions. In order to avoid a familiarity effect the two groups completed the tasks in a different order ((2) and (3)). The FLCAS was administered post-task. Finally, both groups completed the post-experiment questionnaire (4).

Figure 5

The Experiment Procedure Flowchart



Data Analysis

A one-way repeated measures ANOVA was used to compare mean scores for the pre-experiment, and post-task FLA scores. An alpha level of $p = .05$ was set for all statistical tests. The statistical analysis software used was IBM's SPSS 24. A paired samples *t*-test was used to determine if participants' perceptions of the two systems were affected by interactivity.

Research Questions

Based on the study design above, the following questions are explored in this paper:

1. Which system is more effective at reducing learners' FLA?
2. How does interactivity affect learners' perception of task difficulty and enjoyment?

Results

RQ1: FLA Scores

FLA scores measured at the pre-experiment and post-task levels producing three sets of data. Descriptive statistics are available in Table 5. Both the static-VR and interactive-VR systems reduced participants FLA in comparison to their pre-experiment FLA scores.

Table 5

Descriptive Statistics for FLA Scores

FLA score	Mean	Std. Deviation
Pre-experiment	20.04	3.75
Post-task Static-VR	17.13	4.36
Post-task Interactive-VR	16.96	3.90

Following the descriptive statistics, a repeated measures ANOVA was run on the data. Mauchly's test of sphericity indicated that the assumption of sphericity was met and that there was an interaction between FLA scores and when the questionnaire was completed $\chi^2(2) = 4.14$, $p = .11$. Pairwise comparisons revealed that there was a significant difference in FLA mean scores between the pre-experiment and both post-task questionnaires. However, there was no statistically significant difference between mean scores for the two VR modes (see Table 6). This indicates that both systems were effective in lowering participants FLA, but that neither system was more effective than the other.

Table 6

Pairwise Comparisons of Mean FLA Score for All Three Stages of the Experiment

(I) Questionnaire	(J) Questionnaire	Mean Difference (I-J)	Std. Error	Sig.
Pre-experiment	Static-VR	2.917*	0.645	.00
	Interactive-VR	3.083*	0.507	.00
Static-VR	Pre-experiment	–2.917*	0.645	.00
	Interactive-VR	0.167	0.457	1
Interactive-VR	Pre-experiment	–3.083*	0.507	.00
	Static-VR	-0.167	0.457	1

RQ2: Learner Perceptions of the Two Systems

A post-experiment questionnaire explored participant perceptions of learning within each system, with a focus on cognitive demands. Descriptive statistics are provided in Table 7. Of note are two things: 1) all mean scores are higher for the interactive-VR system and 2) all standard deviations are smaller for the interactive-VR. This indicates that converse to predictions, the additional cognitive demands posed by the interactive-VR system were not perceived as such by the participants. A detailed explanation follows.

Table 7

Descriptive Statistics and *t*-test Results for the Post-Experiment Questionnaire Measures

Measure	Mean Scores (*SD*)		Mean diff.	Sig.
	Static-VR	Interactive-VR		
1. It was easy to speak English when using this system.	7.63 (1.59)	8.57 (1.40)	0.93	.002
2. It was fun to learn with this system.	8.77 (1.68)	9.53 (0.9)	0.77	.005
3. It was easy to complete the task with this system.	7.30 (2.22)	8.07 (1.86)	0.77	.032
4. I think this system is an effective way of learning English.	8.40 (1.42)	8.80 (1.19)	0.40	.07
5. I did not feel anxious learning English with this system.	7.57 (2.42)	7.80 (2.21)	0.23	.36

Paired-samples *t*-tests were conducted to compare participants' perceptions of the VR systems for each of the measures on the post-experiment questionnaire. There was a significant difference in the mean scores for three of the five measures. Focusing on measures 1 and 3, results revealed that participants perceived it significantly easier to speak English in the interactive-VR mode ($p < .001$). The interactive-VR task was also considered easier to complete than the static-VR task ($p < .05$). This suggests that the increase in *spatial reasoning* and *steps required* for task completion did not translate into an increase in cognitive task complexity as perceived by participants.

There was also a statistically significant difference in mean scores for measure 2 ($p < .05$) which suggests, again, that increasing the interactive affordances of VR was not perceived as a hindrance but a positive, enjoyable addition. Finally, the differences in mean scores between the systems for measure 4 approached significance ($p = .07$), which suggests that the interactive-VR system was considered a more effective learning environment than the static-VR system. Adding weight to the results of the FLCAS mean scores analysis, there is no significant difference in mean scores for measure 5 here either.

Conclusion

The impetus for the current study was a previous study of ours in which we compared three modes of SCMC on learners' FLA (York et al., 2020). In that study, differences in the three SCMC environments were kept as minimal as possible. This meant that the VR system in the previous study did not utilize the interactive affordances of VR. However, results suggested that compared to oral and video SCMC, the VR environment was the most effective in reducing learners' FLA (York et al., 2020). In the current study, interactive elements were added which were a cause for concern in terms of task complexity. Adding interactive elements to the system was predicted to increase task complexity based on Robinson's Triadic Componential Framework (2011). However, conversely, the results suggest that additional

interactivity created a favorable context for embodied learning which was both enjoyable and easier to communicate than the previous system.

Implications for Designers

For designers, the results of this study suggest that the incorporation of VR's interactive affordances may be beneficial rather than a hindrance for language learning. The provision of feedback (both in terms of interaction with the system, and as indicators of progress) may reduce cognitive demand and increase motivation to learn.

In FTF communication, recasts are considered one of the most frequent types of feedback given by teachers and effective in promoting language development (Long, 2014, p. 27). Within a VR environment, feedback can be given from interlocutors as in a classroom, but also from the environment itself, and based on results of this study, should be included where possible to aid in comprehension.

Implications for Teachers

Assuming the trend in affordable, portable, and stand-alone VR systems continues, it will not be long before devices like the HTC Vive will be implementable in classroom teaching. With such, teachers will be able to offer their students access to immersive environments that can be utilized for language learning purposes: access to native speakers, interactive tasks that take advantage of kinesthetic learning, and embodied co-constructed interaction between peers (see Steffensen, 2015). Results of the current study suggest that not only will learners enjoy using such technology, but that if tasks utilize the affordances of the environment, will be considered less cognitively demanding than tasks done in non-interactive systems.

Future Research

The positive results of this study act as a catalyst for investigating the affective and cognitive benefits of additional VR affordances. From an instructional design perspective, the addition of a time limit and multiple levels of tasks to clear – making the system more like a typical video game – would allow for an individualized experience which matches a learners' developmental stage. This would also keep learners in a flow state (Csikszentmihalyi, 1990). As a concrete example, if a pair of learners could not complete a level within a specific time limit, an easier level would be presented instead. Alternatively, if a pair completed a level in a very short period of time, they could be presented with a much more complex task upon completion (i.e., increasing the number of elements in the spot-the-difference task used here). Finally, as the system is already network-enabled, it is possible to enter the environment from anywhere in the world, and thus, one step that we are keen to explore in future research is how FLA is affected when participants' interlocutors are 1) from another country, or 2) native speakers.

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.

Alizadeh, M. (2019). Augmented/virtual reality promises for ELT practitioners. *JALT2018 – Diversity and Inclusion*, *2018*(1), 360. **https://doi.org/10.37546/jaltpcp2018-48**

Aslan, E., & Ciftci, H. (2019). Synthesizing research on learner perceptions of CMC Use in EFL/ESL writing. *CALICO Journal*, *36*(2), 100–118. **https://doi.org/10.1558/cj.34818**

Bonner, E., & Reinders, H. (2018). Augmented and virtual reality in the language classroom: Practical ideas. *Teaching English with Technology, 18*(3), 33–53.

Cheng, A., Yang, L., & Andersen, E. (2017). Teaching language and culture with a virtual reality game. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. CHI'17: CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery. **https://doi.org/10.1145/3025453.3025857**

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

deHaan, J., Reed, W.M., & Kuwada, K. (2010). The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning & Technology*, *14*(2), 74–94.

Dewaele, J.M., & Thirtle, H. (2009). Why do some young learners drop foreign languages? A focus on learner-internal variables. *International Journal of Bilingual Education and Bilingualism, 12*(6), 635–649. **https://doi.org/10.1080/13670050802549656**

Dickey, M.D. (2005). Three-dimensional virtual worlds and distance learning: Two case studies of active worlds as a medium for distance education. *British Journal of Educational Technology*, *36*(3): 439–451.

Ebuchi, T., & Yokota, Y. (2019). Japan immigration hits record high as foreign talent fills gaps. *Nikkei Asian Review*. **https://asia.nikkei.com/Spotlight/Japan-immigration/Japan-immigration-hits-record-high-as-foreign-talent-fills-gaps**

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

Hawkinson, E., Mehran, P., & Alizadeh, M. (2017). Using MAVR to bring new dimensions to the classroom. *The Language Teacher*, *41*(3), 30–32.

Hewitt, E., & Stephenson, J. (2012). Foreign language anxiety and oral exam performance: A replication of Phillip's MLJ Study. *The Modern Language Journal*, *96*(2): 170–189. **https://doi.org/10.1111/j.1540-4781.2011.01174.x**

Horwitz, E.K., Horwitz, M.B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132.

Iino, A., & Yabuta, Y. (2015). The effects of video SCMC on English proficiency, speaking performance and willingness to communicate. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds.), *Critical CALL – Proceedings of the 2015 EUROCALL conference, Padova, Italy* (pp. 254–260). Research-publishing.net. **http://dx.doi.org/10.14705/rpnet.2015.000342**

Jackson, D.O., & Suethanapornkul, S. (2013). The cognition hypothesis: A synthesis and meta-analysis of research on second language task complexity. *Language Learning*, *63*(2), 330–367. **https://doi.org/10.1111/lang.12008**

Jauregi, K., de Graaff, R., van den Bergh, H., & Kriz, M. (2012). Native/non-native speaker interactions through video-web communication: A clue for enhancing motivation? *Computer Assisted Language Learning*, *25*(1), 1–19. **https://doi.org/10.1080/09588221.2011.582587**

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages, 4(1)*, 13. **http://dx.doi.org/10.3390/languages4010013**

Lim, C.P., Nonis, D., & Hedberg, J. (2006). Gaming in a 3D multiuser virtual environment: Engaging students in science lessons. *British Journal of Educational Technology*, *37*(2), 211–231.

Lin, H. (2014). Establishing an empirical link between computer-mediated communication (CMC) and SLA: A meta-analysis of the research. *Language Learning and Technology*, *18*(3), 120–147.

Liu, D., Dede, C., Huang, R., & Richards, J. (2017). *Virtual, augmented, and mixed realities in education*. Springer.

Long, M. (2014). Second language acquisition and task-based language teaching. John Wiley & Sons.

Melchor-Couto, S. (2017). Foreign language anxiety levels in Second Life oral interaction. *ReCALL*, *29*(1), 99–119.

Melchor-Couto, S. (2018). Virtual world anonymity and foreign language oral interaction. *ReCALL*, *30*(2), 232.

Ministry of Education, Culture, Sports, Science, and Technology. (2014). *On Integrated Reforms in High School and University Education and University Entrance Examination Aimed at Realizing a High School and University Articulation System Appropriate for a New Era*. **https://www.mext.go.jp/component/english/__icsFiles/afieldfile/2015/03/31/1353908_1.pdf**

Paas, F., Tuovinen, J.E., Van Merriënboer, J.J.G., & Darabi, A.A. (2005). A motivational perspective on the relation between mental effort and performance: Optimizing learner involvement in instruction. *Educational Technology Research and Development*, *53*(3), 25–34. **https://doi.org/10.1007/BF02504795**

Reinders, H., & Wattana, S. (2014). Can I Say Something? The effects of digital game play on willingness to communicate. *Language Learning & Technology*, *18*(2), 101–123. **https://doi.org/10125/44372**

Robinson, P. (2000). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, *22*(1), 27–55.
Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *International Review of Applied Linguistics in Language Teaching*, *43*(1), 45–107.
Robinson, P. (2011). Second language task complexity, the Cognition Hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance,* (pp. 3–38). John Benjamins Publishing.
Sasayama, S. (2016). Is a 'complex' task really complex? Validating the assumption of cognitive task complexity. *Modern Language Journal*, *100*, 231–254.
Satar, H.M. & Ozdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal*, *92*(4), 595–613.
Sherman, W.R., & Craig, A.B. (2018). *Understanding virtual reality: Interface, application, and design*. Morgan Kaufmann.
Stein, S. (2019, December 17). *VR in the 2010s: My decade with things on my face*. Cnet.com. **https://www.cnet.com/news/arvr-2009-2019-my-decade-with-things-on-my-face/**
Steffensen, S.V. (2015). Distributed language and dialogism: Notes on non-locality, sense-making and interactivity, *Language Sciences*, *50*, 105–119.
Sweller, J. (2005). Implications of cognitive load theory for multimedia learning. In R.E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 19–30). Cambridge University Press.
Tudini, V. (2007), Negotiation and intercultural learning in Italian native speaker chat rooms. *The Modern Language Journal*, *91*, 577–601. **https://doi.org/10.1111/j.1540-4781.2007.00624.x**
van Gog, T., & Paas, F. (2008). Instructional efficiency: Revisiting the original construct in educational research. *Educational Psychologist*, *43*, 16–26.
Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13*(2), 7–26.
Yanguas, Í., & Flores, A. (2014). Learners' willingness to communicate in face-to-face versus oral computer-mediated communication. *The JALT CALL Journal*, *10*(2), 83–103.
York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2020). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL*, *33*(1), 1–22. **https://doi.org/10.1017/S0958344020000154**
Ziegler, N. (2016). Synchronous computer-mediated communication and interaction. *Studies in Second Language Acquisition*, *38*(3), 553–586. **https://doi.org/10.1017/S027226311500025X**

Appendix A

Spot the difference task: Warm up

In this lesson, you will work with a partner. You will compare two pictures of the same room. There are six items placed in the room, but some of them are in a different place to your partner. This is a spot-the-difference activity. You will have to decide which of the items are the same and which are different.

Activity 1: What is in your living room?

Think of the objects that you find in a living room and make a list of 10 things here:

_____ _____

_____ _____

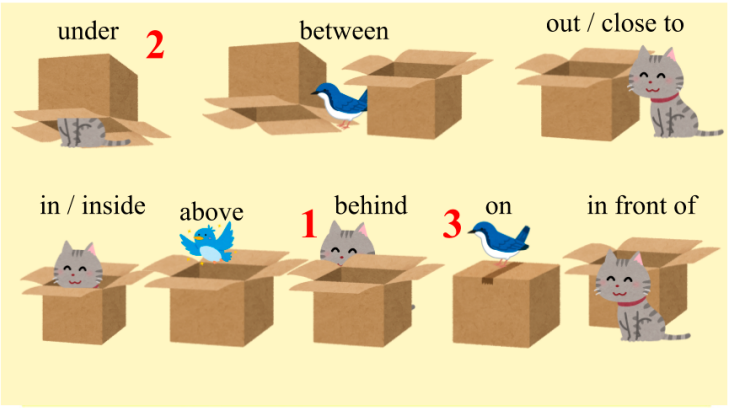
_____ _____

_____ _____

_____ _____

Activity 2: Positions of place

Look at the following picture. The cat 🐱 and bird 🐦 are in different positions in relation to the box.

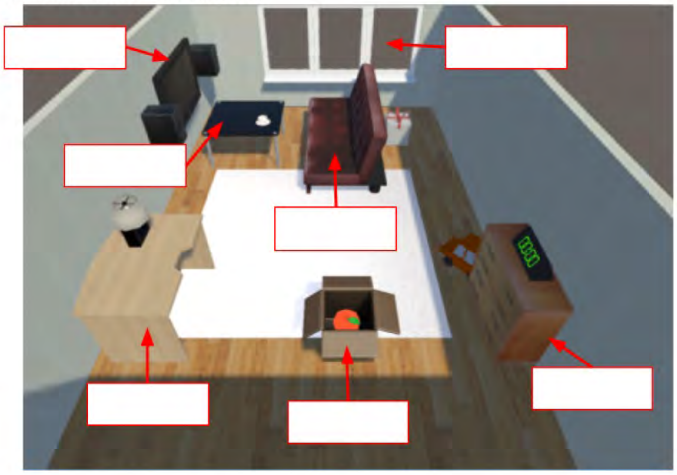


Please write sentences for the three pictures labelled **1**, **2** and **3**. For example: *The cat is in the box.*

1. _____
2. _____
3. _____

Activity 3: Label the items.

Please label the items marked with a textbox:



Activity 4: Where are the objects?

In this room, there are the following objects:

 Apple	 Lamp	 Clock
 Car	 Present	 Cup

Please mark the following sentences correct or incorrect.
Please use a ✓ for correct sentences.
Please use a ✗ for incorrect sentences.
For example:

The apple is on the box. ____X_____

1. The clock is on the chest of drawers. _____
2. The car is in front of the chest of drawers. _____
3. The present is on the sofa. _____
4. The lamp is on the desk. _____
5. The cup is in front of the tv. _____

Authors' Bios

Koichi Shibata is a master's student studying information sciences at Tokyo Denki University. He conducts research on the application of VR in language education.

James York is a lecturer at Tokyo Denki University where he researches the application of games and play in language teaching contexts. He also co-edits *Ludic Language Pedagogy*, an open-access journal.

https://doi.org/10.37546/JALTSIG.CALL2020.9

9

The Effect of Modality on Oral Task Performance in Voice, Video, and VR-based Environments

Hayato Tokutake, James York, and Hiroshi Nakayama, Tokyo Denki University

Abstract

Synchronous computer-mediated communication (SCMC) is a topic of great interest in CALL literature where research has investigated the effectiveness of SCMC compared to traditional face-to-face instruction. However, there are few studies that investigate the intrinsic differences in SCMC modes, particular in terms of their effect on oral communication. At the JALTCALL 2019 conference, we introduced research which assessed the anxiety-reducing affordances of VR. This year we presented results of a follow-up study which focused on the effect of SCMC modality on learners' speaking performance.

30 participants (15 pairs) completed a spot-the-difference task within three different SCMC modes: voice, video, and virtual reality (VR). Using the complexity, accuracy, and fluency (CAF) model, participants' oral task performances were analysed. Results suggest that the voice mode promoted the highest structural complexity, however, the VR mode promoted the highest lexical complexity. Findings therefore suggest that different modes of communication may be used to focus on different skill development. Additionally, practitioners should consider how modality affects learner anxiety and choose the most appropriate system for their students and needs. This paper introduces the VR system, a detailed analysis of results, pedagogical implications, and future research directions for the use of VR in language teaching contexts.

コンピュータを介した同期型コミュニケーション（SCMC）は、CALLの文献で大きな関心が寄せられているトピックであり、従来の対面式の指導と比較してSCMCの有効性が調査されている。しかし、特にオーラルコミュニケーションへの影響の点で、異なるSCMC間での違いを調査する研究はほとんどされていない。JALTCALL 2019では、VRにおける不安軽減アフォーダンスを評価する研究を紹介した。今年は、学習者のスピーキングパフォーマンスに対するSCMCのモダリティの効果に焦点を当てたフォローアップ調査の結果を示した。

本研究では、30名の被験者（15ペア）に、Voice、Video、Virtual Reality（VR）の3つの異なるSCMCで間違い探しのタスクを行わせた。本研究では、複雑さ、正確さ、流暢さ（CAF）モデルを使用して、参加者のスピーキングパフォーマンスを分析した。実験の結果、文法の構造的複雑さにおいて、Voiceモードで最もスコアが良くなった。一方で字句の複雑さにおいては、VRモードが最もスコアが良くなった。

したがって、スキルの開発に焦点を当てる場所によって、異なる通信モードそれぞれに可能性がある

ことが示唆された。そのため、教育者はSCMCのモダリティが学習者の不安にどのように影響するかを検討し、学生の状況に合わせて最適なシステムを選択する必要がある。

この論文では、VRシステム、結果の詳細な分析、教育的影響、および言語教育のコンテキストでVRを使用するための将来の研究の方向性を紹介する。

Keywords: **SCMC, VR, modality, oral proficiency, CAF**

There have always been critiques aimed at the normalization of English as a lingua franca (see Pennycook, 2000), and, entering the third decade of the 21st Century, this issue is as pertinent as ever (Su, 2016). However, the fact remains that speakers of different nationalities may often default to English when communicating with each other. As such, in an increasingly globalized and English-speaking world, the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT) has been promoting the development of practical English skills since 2014. At the primary, secondary, and tertiary levels, there is an emphasis on students being able to "assertively make use of their English skills, think independently, and express themselves" (MEXT, 2014, p. 3).

The spread of English as a Lingua Franca, and intercultural communication in general, may be linked to the rapid development of our networked society and the technology that underpins the vast majority of our daily, increasingly computer-mediated communication. Subsequently, it is not surprising that within foreign and second language studies, there has been a rapid increase in both research and practical implementation of computer-mediated communication (CMC) (Lin, 2014).

This study focuses on the affordances of three different synchronous CMC (SCMC) modalities and their effect on learners' oral task performance. In particular, the study aims to understand what the specific affordances of virtual reality (VR) technology have on learner output as they complete a dialogic task within the VR environment. In order to ascertain whether VR promotes a unique set of communicative affordances, student output is compared with two more frequently studied modalities: voice and video SCMC.

Literature Review

SCMC Studies

CALL research exploring the potential of CMC for language learning has progressed in accordance with technological advances. An early study by Warschauer (1996) explored how telecommunication within a synchronous, text-based CMC domain affected learner motivation and willingness to communicate in comparison to face-to-face (FTF) communication. Findings suggested that students participated in conversations more equally in the CMC mode with an increase in lexical complexity and formality. This study therefore provided empirical evidence that modality does indeed affect learner output. As a critique of the study, and related to the present study, it is not surprising that there was a difference in lexical complexity for both modes, as this difference exists between written and oral output before even considering CMC.

At the time of Warschauer's study, synchronous CMC modalities that incorporated voice

and video were not in widespread use. However, since then, such studies (i.e., those that compared voice and video-based CMC with FTF communication) have started to appear. Voice and video modes are more similar to FTF in that they primarily rely on oral communication. The word *primarily* is used in the previous sentence because the majority of voice and video-based CMC tools emerged as an extension of preexisting text-based modes. For example, although Skype is a VOIP-based voice and video communication tool primarily, it is multimodal in that it also supports text-based interactions. One particular study which acted as the impetus for a previous study of our own (York et al., 2021) was that of Satar and Özdener (2008). They investigated the effects of SCMC on speaking proficiency and anxiety. Two modes were employed: text and voice SCMC. Sixty participants completed language activities in one of the two modes in addition to their usual classroom-based instruction. A control group of 30 students did not complete any of the additional activities, thus a total of 90 participants were used in the study. Based on speaking proficiency tests at the end of the treatment, it was discovered that the speaking proficiency of both of the experimental groups increased in comparison to the control group in a post-experiment speaking test. Thus, one important finding of this study was that the text-based SCMC group were able to improve their speaking proficiency without physically practising speaking. However, in comparison with the present study, Satar and Özdener's study lacked a rigorous evaluation of learner output *during* task performance, only their ability to perform during a post-task test.

Following on from Satar and Özdener's study, Yanguas (2010) added video-based SCMC as a modality in a study which explored how the addition of video affected learner output. Fifteen learner dyads (i.e. pairs) were created and assigned to complete an oral communication task within a voice-based, video-based, or face-to-face modality. The study assessed learner output performance in terms of the number of times they negotiated for meaning (NfM) with their partner. Findings suggested that there was no difference in terms of NfM between the video-based SCMC mode and the FTF mode, however, there was an increase in NfM for the dyads that completed the task in the voice-based SCMC mode. The reason for this difference was attributed to the lack of visual stimulus, meaning that the dyads had to rely on verbal explanations instead.

In addition to the tools outlined above (voice, video, and text based SCMC), some studies have investigated the affordances of SCMC within virtual worlds (VWs), and digital games, which are becoming increasingly sophisticated (Melchor-Couto, 2017; 2018; Wigham & Chanier, 2015). York (2019) operationalized tasks for one such virtual world: *Minecraft*. In a counterbalanced, repeated-measures study, 10 dyads carried out three sets of oral communication tasks in two different environments: within *Minecraft* and FTF (totalling 6 tasks). Participants' oral proficiency was analysed in terms of output complexity, accuracy and fluency, and it was found that the VW hindered fluent output, but that modality had no significant effect on accuracy and complexity dimensions. Instead, task *type* appeared to be a more influential factor. Results echoed a similar study by Yanguas and Bergin (2018), who found no difference in the number of learning related episodes (LREs) when participants completed jigsaw tasks in video and voice based SCMC modes.

Finally, related to the use of virtual worlds as domains for CMC, there are a number of meta-analyses on the use of SCMC in language learning and teaching environments.

Focusing on one recent analysis, Zeigler (2016) analysed the results of 14 studies which compared the effectiveness of SCMC to that of FTF communication. Studies were selected based on their inclusion of interaction between participants during task performance. Findings of the meta-analysis found that overall, there was no significant effect between SCMC and FTF on learner interaction. Notably, of the 14 studies chosen, none featured the use of a virtual domain and only three implemented oral SCMC as the modality of focus. This highlights a lack of focus and a need for more studies which operationalize tasks within an oral SCMC mode. The present study hopes to fill this gap by utilizing three different SCMC oral modes.

VR and Second Language Interaction

VR and its application in language learning is a nascent field, with few empirical studies to date. The majority of research papers on the use of VR are hypothetical and based on the possible affordances of the technology in language learning and teaching contexts. Hawkinson et al. (2017) introduced several different reality-augmenting technologies, differentiating between augmented reality, virtual reality, and mixed reality modes. Following, they connected these technologies to cognitive science studies, introducing the potential of such tools for language learning. Similarly, Bonner and Reinders (2018) introduced a number of ways that VR could be utilized in classroom teaching, which is a first step to its implementation. Their ideas included virtual campus or city tours, as well as practicing public speaking via specific apps.

Regarding empirical studies, York et al. (2021) researched the affective affordances of VR as a domain for second language interaction. Their study assessed the effect of three different SCMC modes on learners' foreign language anxiety (FLA): voice, video, and VR. Results of quantitative data analysis revealed that all three modes reduced learners FLA compared to pre-test FLA scores, however, learner perceptions suggested that of the three environments VR was considered the most effective domain for language learning, as well as being the easiest to communicate with interlocutors. Findings therefore align somewhat with those of Melchor-Couto (2017) who found that learner FLA was reduced when learning within the virtual world Second Life compared to an FTF group. The reduction in FLA when communicating in virtual domains has been attributed to the *avatar effect* where the avatar provides users with an anonymous shield through which to talk (Lotherington & Jenson, 2011). Subsequently, Xie et al. (2019) explored the affordances of a VR system on oral proficiency. Two groups, one with the aid of a VR environment and one without, performed a monologic, oral task. Results suggested that both the content and vocabulary of participants' oral were influenced by mode. The VR group had statistically significantly higher scores than the group without the aid of the VR environment, and the researchers concluded that access to additional stimuli may have influenced the higher scores. In summary, research on VR has typically proposed practical ways in which it could be implemented in classroom contexts, though there are few empirical studies appearing in the literature.

Research Question

The aim of the current research is to uncover the effect of SCMC modes on oral task performance. As such, the research question is:

1. Is there a difference in learner oral proficiency when completing interactive tasks within voice, video and VR-based SCMC modes?

As a follow up, sub-question: If there is a difference in performance based on modality, what are the affordances for each domain for promoting language proficiency?

Methodology

Participants

The study was undertaken at a private science and engineering university in Saitama, Japan. Thirty participants volunteered to take part in the study. Their mean age was 20.9 (*SD* = 1.12). They were all native speakers of Japanese and had received the same amount of formal English education, thus making this a homogeneous group of learners (see Table 1). All participants provided their informed consent via a consent form at the start of the study. The form outlined the study purpose and data handling procedures including information regarding the anonymization of data.

Table 1

Baseline Characteristics of Participants

Baseline Characteristic		n	%
Age			
	19	2	6.7
	20	11	36.7
	21	9	30.0
	22	6	20.0
	23	1	3.3
	24	1	3.3
Gender			
	Male	26	86.7
	Female	4	13.3
Native Language			
	Japanese	30	100

In order to avoid a familiarity effect, the researchers created dyads by pairing participants with partners that they did not know and met for the first time on the day of the study.

Dyads completed a spot-the-difference task in all three domains, and so a counterbalance design was used to prevent the resultative effect in which task familiarity increases a student's performance and confidence in their performance (see Skehan, 1991).

Additionally, dyads completed the tasks in different orders in a counterbalanced, repeated-measures design. Three groups were created, and are referred to as Group A, B, and C.

Instruments

The three SCMC modes utilized in this study were oral (henceforth voice), video, and VR (developed in Unity **https://unity.com/**) A spot-the-difference task was created in the VR domain and the same assets were reused in both the voice and video and SCMC modes in order to unify the content for each task.

VR modality

Figure 1

An Example of the Spot-the-Difference Task in VR (Participant 1)

Figure 2

An Example of the Spot-the-Difference Task in VR (Participant 2)

As can be seen in Figures 1 and 2, the spot-the-difference task featured a doll's house that each participant could look into, as well as seeing their interlocutor (Figure 3). The number of differences in each task was operationalized to three per task. Avatars were utilized as embodied versions of the participants. These avatars echoed participants' body, head, and arm movement. This was achieved with the use of the HTC Vive head mounted display unity (see Figure 4).

Figure 3

A Screenshot of the VR Task Showing the Interlocutor

Figure 4

A Participant Wearing the HMD During the VR Task

Voice and Video Modalities

For the voice and video modalities, the position of items within the VR doll houses were changed and screenshots taken to create different versions of the spot-the-different task.

These screenshots were presented to learners on a monitor screen (see Figure 5). The difference between the voice and video modalities was whether interlocutors were visible or not (see Figure 5 and Figure 6).

Figure 5

A Participant Completing the Spot-the-Difference Task in the Voice Modality

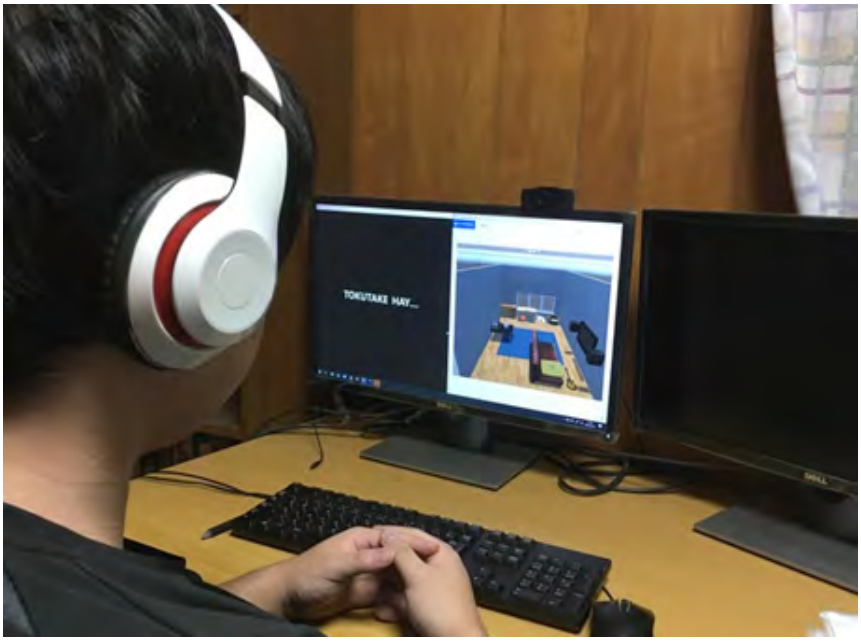


Figure 6

A Participant Completing the Spot-the-Difference Task in the Video Modality

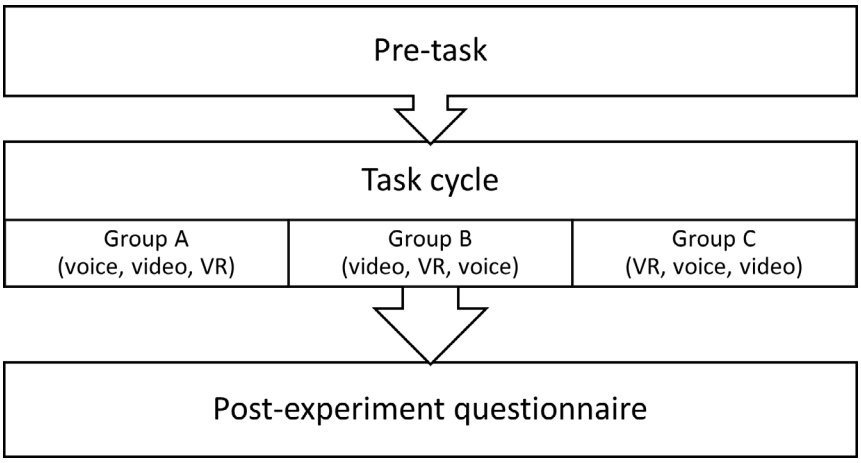
Zoom was employed as the communication software for all tasks, allowing for the recording of both audio and video.

Procedure

The experiment was a repeated measures design, where each pair of participants completed the spot-the-difference task in all three modes. The design is therefore based on that of York et al., (2020). Three groups of pairs (10 pairs per group) were created which completed tasks in a different order to avoid the Resultative Hypothesis (Skehan, 1991). That is, the mastery of tasks can affect learners' motivational attitudes and performance, and so the modality of task repetitions varied between groups (see Figure 7).

Figure 7

Experiment Flowchart



Before undertaking the spot-the-different tasks in each of the three SCMC modes, all participants completed a pre-task to prime them for the main task (see Appendix A). Upon completion of the pre-task, pairs carried out the spot-the-difference tasks in the order assigned to their group. The task cycle included an explanation of the task and instructions on how to complete the task within that domain (for instance, the VR task included a short demonstration on how to use the controllers to gesture to their partner).

Upon completing all three tasks participants answered questions on a post-experiment questionnaire. The questionnaire was devised to measure foreign language anxiety and collect perceptions of learning within each modality. For this paper, the only question that is referenced is the final open-ended question which allowed participants to comment on their experiences. This data was referenced as participant perceptions and used to provide further insight into any findings.

Data Analysis: The CAF Model

The CAF model is a robust way of assessing the proficiency of learner output along multiple, quantitative dimensions (Housen & Kuiken, 2009). The model examines learner output in terms of complexity, accuracy, and fluency (henceforth, CAF). The genesis of the model is associated with the work of Skehan (1998). There is a plethora of measures available in the literature for assessing learner output and so the following three sections will introduce those used in this study, including a rationale for their inclusion.

Complexity Measures

Learner output complexity can be measured in two ways: Structural complexity and lexical complexity. In this study, two methods for assessing output complexity were utilized, one for each complexity type.

- **Structural complexity** was measured by counting the number of words per utterance (Ortega, 1999).
- **Lexical complexity** was measured by the total number of different words a participant produced (number of types, not tokens).

The website Lextutor (**https://www.lextutor.ca/**) was utilized in accessing the number of types.

Accuracy Measure

Accuracy was measured by assessing the number of error-free utterances a participant produced. This measure resembles the most common measure found in the literature: error-free clauses (Skehan & Foster, 1997; Yuan & Ellis, 2003; Bygate, 2001). However, it was modified here due to the nature of participant speech. Participants did not produce utterances that contained multiple clauses (see also York, 2019).

Fluency Measures

There are two major measures for fluency: **temporal** and **vocal** fluency. Temporal fluency refers to a participant's rate of speech and is typically measured as words or syllables per minute (Yuan & Ellis, 2003). Vocal fluency refers to how "exact" a participant speaks, thus focusing on pauses, repetitions, and number of false starts (Skehan, 2009). In this study, participant fluency was measured as the number of words spoken per minute.

Statistical Tests

To prepare data for statistical tests, pairs were recorded and later the audio transcribed by the authors. The transcriptions were then coded for accuracy. Codes used were *Y* for an error-free utterance and *N* for an erroneous utterance. A sample can be seen in Table 2.

Table 2

Sample Transcription

Utterance Number	Participant Number	Utterance	Accuracy Codes
47	1	It is a difference point.	y
48	2	Yes	y
49	1	The guitar in front of sofa	n
50	2	Yes	y
51	1	OK	y
52	2	Where is the chest?	y
53	1	Chest is the desk left side.	n
54	2	Yes.	y
55	1	The same.	y
56	2	The same.	y
57	1	What color is carpet?	n

One-way repeated measures ANOVA analyses were utilized to look for statistically significant differences in mean scores for all measures. SPSS (version 24) was used to conduct analyses. Subsequently, a post hoc Bonferroni adjustment was used to understand which modalities' mean scores differed significantly. The alpha level for all statistical tests was set at $p < .05$.

Results

First, a detailed examination of participants' output in terms of complexity, accuracy, and fluency is presented. Based on any findings, the affordances for each mode are then explored.

Complexity

Mean scores for the number of **words per utterance** measure are displayed in Table 3. The highest mean score was recorded for the voice modality (2.82), the lowest was recorded for the VR modality (2.56).

Table 3

Mean Scores for the Number of Words per Utterance Measure

Modality	Mean	*SD*
Voice	2.82	0.473
Video	2.68	0.55
VR	2.56	0.398

Inspection of the results of the one-way repeated measures ANOVA revealed that there was a statistically significant interaction ($p < .05$), and pairwise comparisons revealed that there was a statistically significant difference between mean scores for the Voice and VR modes (Table 4 and Figure 8).

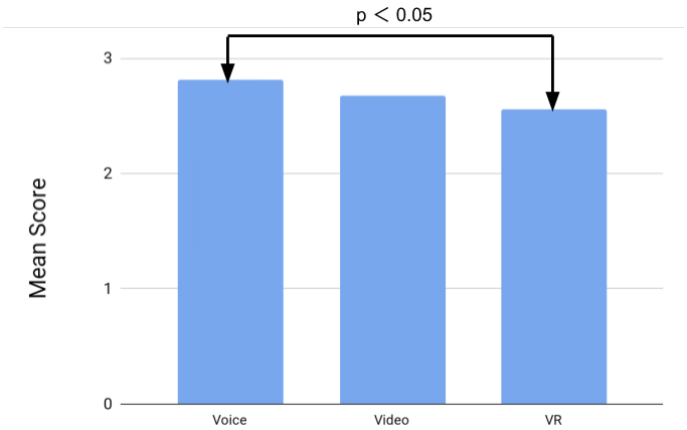
Table 4

Pairwise Comparisons for Simple Main Effects of Modality on Words per Utterance

Modality (I)	Modality (J)	Mean difference(I-J)	Sig
Voice	Video	0.14	0.23
Voice	VR	0.26	0.01*
Video	VR	0.12	0.18

Figure 8

Graphical Representation of Mean Scores for Number of Words per Utterance Measure



Regarding the second complexity measure: **number of different words spoken**, mean scores for each modality are presented in Table 5. The highest mean score was recorded for the VR mode, and the lowest for the video mode.

Table 5

Mean Scores for the Number of Different Words Measure

Modality	Mean Score	*SD*
Voice	31.23	9.62
Video	28.20	11.17
VR	36.97	17.25

Inspection of the results of the one-way repeated measures ANOVA revealed that there was a statistically significant interaction ($p < .01$), and so pairwise comparisons were conducted. Results showed a statistically significant difference between the video and VR modes (Table 6 and Figure 9).

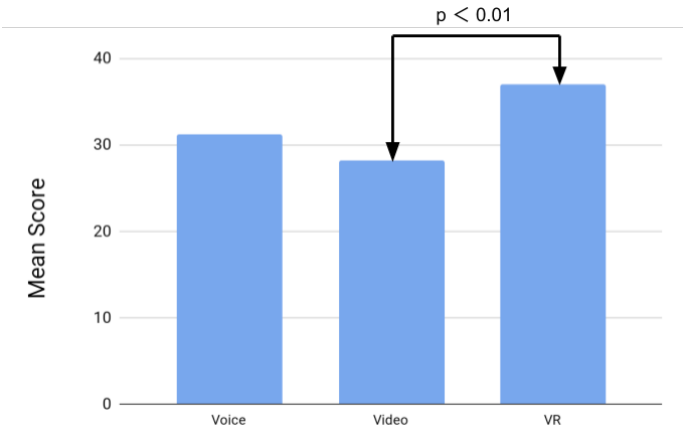
Table 6

Pairwise Comparisons for Simple Main Effects of Modality on the Number of Different Words Produced

Modality (I)	Modality (J)	Mean difference(I-J)	Sig
Voice	Video	3.03	0.36
Voice	VR	5.74	0.09
Video	VR	8.77	0.00**

Figure 9

Graphical Representation of Mean Scores for Different Words Measure



Accuracy

Table 7 shows the mean scores for participants error free utterances for each modality. Participants produced the most error-free utterances when completing the spot-the-difference task in the video mode, though all mean scores are similar. Inspection of the results of the one-way repeated measures ANOVA revealed no statistically significant interaction between mean scores.

Table 7

Mean Scores for Error-free Utterances Measure

Modality	Mean Scores	*SD*
Voice Chat	0.64	0.215
Video Chat	0.70	0.165
VR Chat	0.67	0.184

Fluency

Mean scores for the fluency measure employed in this study: **number of words per minute** are presented in Table 8. Though scores for each mode are similar, the highest mean

score was recorded for the voice mode, and the lowest for the VR mode. Inspection of the results of the one-way repeated measures ANOVA revealed no statistically significant interaction.

Table 8

Mean Scores for Participants' Number of Words per Minute

Modality	Mean Score	*SD*
Voice Chat	29.46	15.41
Video Chat	28.75	15.413
VR Chat	28.05	8.11

Post Experiment Questionnaires

Of the 30 participants, 26 gave comments to the open-ended question asking for feedback on their experiences. There were 18 comments in relation to the VR mode, 7 for video, and 2 comments recorded for the oral mode. A selection of typical comments for each mode are provided below.

A number of comments in relation to the VR mode mentioned that it was an effective domain for language learning, it was interesting, and the embodied experience was a catalyst for easier communication than other modes (see Table 9 for representative comments).

Table 9

Positive Comments Related to the VR Mode

Participant Number	Comment (translated from Japanese to English)
1	This was my first-time using **VR**, and I found it incredibly interesting.
2	I was able to see my partner's gestures when using the **VR** system, and so it felt like I was talking to somebody in front of me. It made me feel relaxed.
6	Speaking to someone in English using **video** or **VR** was a new experience for me and so I completed the tasks with a feeling of intrigue.
14	I was surprised that I didn't feel seasickness when completing the **VR** task.
16	**VR** was the most fun. I'm looking forward to a practical implementation of this system.
20	It was easier to complete the tasks with the **video** and **VR** systems because we could see our partner. With **voice** only, I didn't know when my partner was ready.
23	When using **video** the task screen and partner screen were separate, so it was easier to communicate in the **VR** domain because it was just one context.
29	I think **VR** will be very appealing to students because it is a currently hyped media. For that reason, I think it is an effective learning environment.

However, several negative comments were recorded for the VR domain (Table 10). These were related to the cognitive demands placed on learners due to the novelty of the

environment and the ease of gesturing in the domain meaning that they did not need to rely on their language skills. One comment mentioned that the VR mode placed increased spatial reasoning demands on participants, making it a more difficult domain to complete the task (Participant 9).

Table 10

Negative Comments Related to the VR Mode

Participant Number	Comment (translated from Japanese to English)
5	I was a little puzzled when using **VR** because I wasn't used to it.
9	It was hard to understand my partner's directions in **VR** because of the more complex spatial properties of the system
23	I was not very familiar with **VR** and so I didn't use many gestures. It was easier to gesture when doing the **video** task.
24	Because my partner was not directly in front of me during the **VR** task but off to the side, it was a little difficult to look over and check their gesture.
28	Because it was easier to gesture in **VR**, we didn't have to rely on our English ability so much..

Regarding the video mode, three comments stood out as negative opinions regarding this environment (Table 11). Comments generally mentioned that participants were more focused on their own image or doll's house and that they did not make use of the affordances of the video of their partner. Additionally, one comment mentioned feeling embarrassed by being visible to their partner.

Table 11

Comments Regarding the Video Mode

Participant Number	Comment (translated from Japanese to English)
12	Regarding the **video** and **voice** systems, although you could see your partner when using video, I spent more time looking at the picture, so I didn't look at my partner's face very much. I felt no difference between these two systems.
13	Speaking honestly, I didn't notice too much difference between the **VR** and **video** environments because I focused on my room rather than my partner.
23	During the **video** task, I couldn't look at my partner. I was also a little embarrassed to have my own **video** showing.

Finally, comments regarding the voice mode revealed that, due to lack of visual clues, one participant did not know when their partner was ready (Participant 20) and that there was no major difference between the video and voice modes (Participant 12).

Discussion

The aim of the present study was to assess the effect of three SCMC domains on learner oral output as they completed a dialogic task in each environment. Statistical tests revealed that there was a statistically significant difference in participants' output complexity along both measured dimensions. First, regarding the structural complexity measure (i.e., number of words per utterance participants produced between the voice and VR modes), participants produced more words per utterance when completing the voice task. As seen in Yanguas (2010), the lack of visual clues (gestures and other paralinguistic clues) when completing the task in the voice mode could have forced students to rely on their spoken language. Similarly, as seen in York (2019), the cognitive demands of the VR mode could have hindered complex output. Referring to comments, cognitive demands could have been increased due to a lack of familiarity with completing tasks in VR environments, and the use of gestures could have lessened the need for speaking compared to other modes.

The lexical complexity measure (i.e., number of different words spoken) also revealed a statistically significant difference between the video and VR modes. Participants produced a greater volume of different words when completing the VR task in comparison to the video mode. Reasons for this were not immediately evident in participant comments, however, as seen in Xie et al. (2019), the additional stimuli presented by the VR environment could have allowed for an increased awareness of and access to vocabulary.

As for accuracy and fluency measures, no statistically significant differences were found in the data. This echoes York (2019), who also found that task type was more influential than modality on learners' output accuracy and fluency.

Regarding the sub-question "What are the affordances of each domain for language learning," a simple summary follows. Results suggest that the voice mode may be a suitable domain for promoting student fluency, and the VR domain for focusing on vocabulary learning.

Conclusion

This study aimed to uncover the cognitive affordances of three unique SCMC environments. The specific research question was, "Is there a difference in learner oral proficiency when completing interactive tasks within voice, video, and VR-based SCMC modes?" Thirty students (15 dyads) completed a spot-the-difference task in all three modes where their oral proficiency during task performance was measured along complexity, accuracy, and fluency dimensions. Results of this study suggest that the VR mode pushed learners to produce the highest lexical complexity, however, for other measures, no statistically significant differences were found. This suggests that all three modes may be effective in promoting oral proficiency. A negative interpretation of this would be that if the same results can be achieved with simple voice-SCMC, that is, without spending excessive money on VR environments, then teachers have no reason to worry themselves with keeping up with EdTech trends and demand VR systems for their students. However, with the increasing complexity and possibilities that VR environments provide, there is a great potential of this modality to support language learning in future classrooms. The current study utilized very

little of VR's affordances (interactivity, feedback, access to native speakers, etc.) and so our own future studies will focus on incorporating more features of VR and assessing how such features affect language learning from both cognitive and affective perspectives.

The implications of the current study for teachers is that the embodied experience of completing tasks in VR may be both a source of increased engagement, but due to reduced FLA, may also be beneficial in improving learners' oral proficiency. Of course, we are aware that the head-mounted display supported VR environment used in this project is far from being practically implementable in language classrooms, but similar VR systems are being developed and rapidly becoming more portable (see the Oculus Quest for example of a standalone system). Future research should increase the number of task types that participants complete to explore how the VR modality affects performance along the task complexity dimension (see York, 2019). The effect of different VR systems on learner output should also be explored to understand if graphical fidelity and interactivity is an influential component of VR (e.g., smartphone VR, standalone VR sets, desktop PC-enabled VR, etc.).

References

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Longman.

Bonner, E., & Reinders, H. (2018). Augmented and virtual reality in the language classroom: Practical Ideas. *Teaching English with Technology*, *18*(3), 33–53.

Hawkinson, E., Mehran, P., & Alizadeh, M. (2017). Using MAVR to bring new dimensions to the classroom. *The Language Teacher*, *41*(3), 30–32.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 461–473. **https://doi.org/10.1093/applin/amp048**

Lin, H. (2014). Establishing an empirical link between computer-mediated communication (CMC) and SLA: A meta-analysis of the research. *Language Learning & Technology, 18*, 120–147.

Lotherington, H., & Jenson, J. (2011). Teaching multimodal and digital literacy in L2 settings: New literacies, new basics, new pedagogies. *Annual Review of Applied Linguistics*, *31*, 226–246.

Melchor-Couto, S. (2017) Foreign language anxiety levels in Second Life oral interaction. *ReCALL, 29*(1): 99–119. **https://doi.org/10.1017/S0958344016000185**

Melchor-Couto, S. (2018) Virtual world anonymity and foreign language oral interaction. *ReCALL, 30*(2): 232–249. **https://doi.org/10.1017/s0958344017000398**

Ministry of Education, Culture, Sports, Science, and Technology. (2015). On integrated reforms in high school and university education and university entrance examination aimed at realizing a high school and university articulation system appropriate for a new era. **https://www.mext.go.jp/component/english/__icsFiles/afieldfile/2015/03/31/1353908_1.pdf**

Ortega, L. (1999) Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, *21,* 109–148.
Pennycook, A. (2000). English, politics, ideology: From colonial celebration to post-colonial performity. In T. Ricento (Ed.), *Ideology, politics, and language politics: Focus on English* (pp. 107–120). John Benjamins.
Satar, H.M. & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal*, *92*(4), 595–613.
Skehan, P. (1991). Individual differences in second language learning. *Studies in Second Language Acquisition*, *13*(2), pp. 275–298.
Skehan, P. (1998). *A cognitive approach to language learning.* Oxford University Press.
Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, *30*(4), 510–532. **https://doi.org/10.1093/applin/amp047**
Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task based learning. *Language Teaching Research, 1*(3), 185–211.
Su, Y.C. (2016). The international status of English for intercultural understanding in Taiwan's high school EFL textbooks. *Asia Pacific Journal of Education, 36*(3), 390–408. **https://doi.org/10.1080/02188791.2014.959469**
Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal, 13*(2), 7–26.
Wigham, C.R., & Chanier, T. (2015) Interactions between text chat and audio modalities for L2 communication and feedback in the synthetic world Second Life. *Computer Assisted Language Learning*, *28*(3): 260–283. **https://doi.org/10.1080/09588221.2013.851702**
Xie, Y., Chen, Y., & Ryder, L.H. (2019). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning*, 1–21. **https://doi.org/10.1080/09588221.2019.1604551**
Yanguas, I. (2010). Oral computer-mediated interaction between L2 learners: It's about time! *Language Learning & Technology, 14*(3), 72–93.
Yanguas, I., & Bergin, T. (2018). Focus on form in task-based L2 oral computer-mediated communication. *Language Learning & Technology, 22*(3), 65–81. **https://doi.org/10125/44657/**
York, J. (2019) Language learning in complex virtual worlds: Effects of modality and task complexity on oral performance between virtual world and face-to-face tasks. [Doctoral dissertation, University of Leicester]. **https://leicester.figshare.com/account/articles/10242452**
York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2021). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL*, *33*(1), 49–70. **https://doi.org/10.1017/S0958344020000154**

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**

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction. *Studies in Second Language Acquisition, 38*(3), pp. 553–586. **http://doi.org/10.1017/S027226311500025X**

Appendix A

Pre-Task worksheet

Spot the difference task: Warm up

In this lesson, you will work with a partner. You will compare two pictures of the same room. There are six items placed in the room, but some of them are in a different place to your partner. This is a spot-the-difference activity. You will have to decide which of the items are the same and which are different.

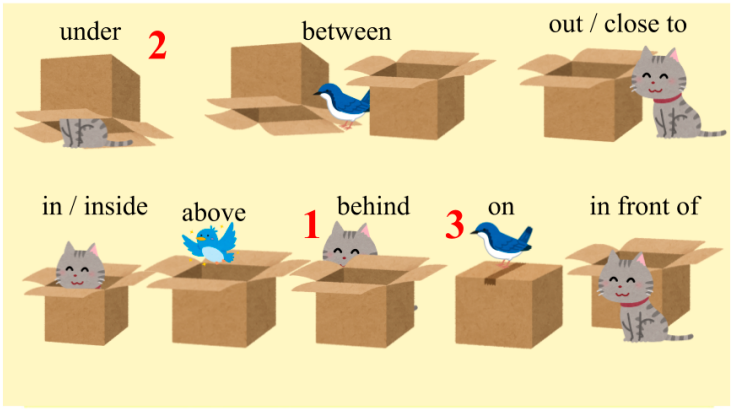
Activity 1: What is in your living room?

Think of the objects that you find in a living room and make a list of 10 things here:

_____	_____
_____	_____
_____	_____
_____	_____
_____	_____

Activity 2: Positions of place

Look at the following picture. The cat 🐱 and bird 🐦 are in different positions in relation to the box.

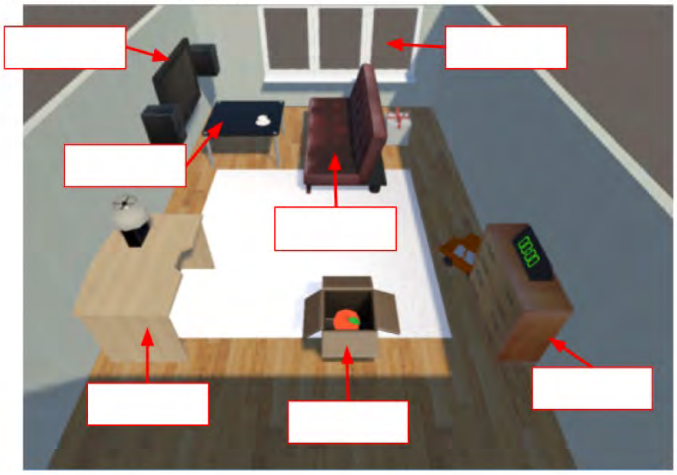


Please write sentences for the three pictures labelled **1**, **2** and **3**. For example: *The cat is in the box.*

1. _____
2. _____
3. _____

Activity 3: Label the items.

Please label the items marked with a textbox:



Activity 4: Where are the objects?

In this room, there are the following objects:

 Apple	 Lamp	 Clock
 Car	 Present	 Cup

Please mark the following sentences correct or incorrect.
Please use a ✓ for correct sentences.
Please use a ✗ for incorrect sentences.
For example:

The apple is on the box. ____X_____

1. The clock is on the chest of drawers. _____
2. The car is in front of the chest of drawers. _____
3. The present is on the sofa. _____
4. The lamp is on the desk. _____
5. The cup is in front of the tv. _____

Authors' Bios

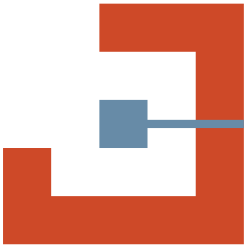
Hayato Tokutake is a Master's student studying information sciences at Tokyo Denki University. He conducts research on the application of virtual reality in language education.

James York is a lecturer at Tokyo Denki University where he researches the application of games and play for teaching languages. He also co-edits *Ludic Language Pedagogy,* an open-access journal.

Hiroshi Nakayama is a professor of instructional technology at Tokyo Denki University, conducting research on the integration and application of technology for educational purposes.

ISBN 978-4-901352-63-5

Teaching with Technology 2020
Selected papers from the JALTCALL2020 Conference



JALT CALL SIG **https:// jaltcall.org**