

Proceedings

of The XXIst International CALL Research Conference

Edited by Jozef Colpaert Yijen Wang Glenn Stockwell



2202



Licensed to Pham Vu Phi Ho (Email: ho.pvp@vlu.edu.vn/Tel: 84909850699) on July 13, 2022. Copying or distribution strictly prohibited by law.

Licensed to Pham Vu Phi Ho (Email: ho.pvp@vlu.edu.vn/Tel: 84909850699) on July 13, 2022. Copying or distribution strictly prohibited by law.

Proceedings of the XXIst International CALL Research Conference

Edited by

Jozef Colpaert University of Antwerp

Yijen Wang Waseda University

Glenn Stockwell

Waseda University



Melbourne – London – Tokyo – New York



www.castledown.com

4th Floor, Silverstream House, 45 Fitzroy Street Fitzrovia, London W1T 6EB United Kingdom

Ground Level, 470 St Kilda Road, Melbourne, Victoria 3004, Australia

447 Broadway, 2nd Floor #393, New York NY, 10013 United States

2nd Floor Daiya Building, 2-2-15 Hamamatsu-cho, Minato-ku, Tokyo 105-0013, Japan

First published 2022 by Castledown Publishers, London

Information on this title: www.castledown.com/academic-books/view-title/?reference=9781914291050 DOI: 10.29140/9781914291050

Proceedings of the XXIst International CALL Research Conference

© Jozef Colpaert, Yijen Wang & Glenn Stockwell (Editors), 2022

All rights reserved. This publication is copyright. Subject to statutory exception and to the provisions of relevant collective licencing agreements, no reproduction, transmission, or storage of any part of this publication by any means, electronic, mechanical, photocopying, recording or otherwise may take place without prior written permission from the author.

Typeset by Castledown Design, Melbourne

ISBN: 978-1-914291-05-0 (Paperback) ISBN: 978-1-914291-06-7 (Digital)

Castledown Publishers takes no responsibility for the accuracy of URLs for external or third-party internet websites referred to in this publication. No responsibility is taken for the accuracy or appropriateness of information found in any of these websites.

Table of Contents

Acknowledgements	viii
Actualizing the affordances of Machine Translation Tools for language learning Antonie Alm	1
Uncovering the role of learning ecology in explaining students' engagement in informal L2 learning activities in digital online environments Artem Zadorozhnyy, Yu Baohua	7
The role of technology and technology training in language teachers' professional development in the private sector Bao Nguyen	16
Using screencasts to create personalized formative feedback in academic writing courses Bradley Irwin	20
Comparing the effectiveness of online and in-class collaborative writing Cao Thi Xuan Tu, Pham Vu Phi Ho	23
A corpus of short YouTube news videos to inform course design and materials development in an EFL university setting in Japan Christopher Robert Cooper	33
Repetition supports the effects of Involvement Load Hypothesis on improving students' productive vocabulary performance Guoyuhui Huang	47
A pilot study of children with dyslexia and learning foreign languages using 3D letters Hanae Ikeshita	55
EFL learners' oral communication strategies: Insights from a Taiwan and Poland exchange project Hsin-chou Huang, Barbara Loranc	64
Applying translanguaging pedagogy to scaffold non-English major juniors on writing scripts for English presentations Hsin-Yi Cyndi Huang, Chiung-Jung Tseng, Ming-Fen Lo	72
A qualitative study on using Intelligent Personal Assistant in teaching a young Korean learner in an EFL context Hyangeun Ji, Soyeon Park, Jung Won Jo	78
The Matthew effect in CALL: Examining the equity of a novel intelligent writing assistant as English language support John Maurice Gayed, May Kristine Jonson Carlon, Jeffrey Scott Cross	80

Theoretical foundation in designing and developing a mobile app to support ESL learning for STEM learners Karmila Rafiqah M. Rafiq, Harwati Hashim, Melor Md Yunus			
Culture, confidence and connections: Telecollaboration as a springboard for successful JFL learning Mark R. Freiermuth, Ikuko Tomida	99		
Multimodal humor in emoji-mediated emotive communication Mei-Ya Liang	107		
Computer mediated communication and task-based learning for adolescent learners of Chinese as a foreign Language in Ireland: An eBook task design under the adaption of Bridge 21 technology-mediated learning model Mengdi Wang, Ann Devitt, Juan Gao, Ciarán Baurer	117		
Digital game-based SLA in the wild: evidence from a qualitative case study Michael Hofmeyr	126		
Using Socrative to facilitate problem-based learning in large undergraduate courses Myung Jeong Ha	132		
A mobile diary application as an instrument for collecting real-world and real-life contextualized language learning Orit Ezra, Anat Cohen, Dror Shany	137		
On the adequacy of L2 pronunciation feedback from automatic speech recognition: A focus on Google Translate Paul John, Walcir Cardoso, Carol Johnson	146		
Effect of English language proficiency on learner engagement in communities of inquiry Phuong Tran	155		
Pedagogical benefits of technological affordances in a user- created metaverse space Sangmin-Michelle Lee	159		
Exploring the impact of AI on EFL teaching in Japan Stephen Lambacher, Hisayo Kikuchi, Hiroyuki Obari	163		
Towards truly intelligent and personalized ICALL systems using Fluid Construction Grammar Veronica Juliana Schmalz, Frederik Cornillie	169		
Computer-mediated communication (CMC) text analytics: Exploring the dynamics within digital discourse Ward Peeters	180		

Investigating college students' perceptions of online and offline review modes in academic writing courses Yi-Chin Hsieh, Alvin Ping Leong				
A study on social interactions among primary students in English vocabulary acquisition in a mobile learner-generated content learning environment Yin Yang, Yanjie Song	194			
Examining primary student self-regulated vocabulary learning behavioural patterns and vocabulary learning outcomes leveraged by the mobile app with a self-regulation scheme Yin Yang, Yanjie Song	204			
A two-year investigation of a Facebook community for supporting	212			

Acknowledgements

We are delighted that the XXIst International Research Conference has finally come to fruition despite the multiple burdens that we faced along the way in trying to host it. Waseda University was to host the conference in 2020 for the first time in Japan, and the conference team was very much looking forward to welcoming visitors from around Japan and the world to participate. Then in February 2020, the world was struck with the COVID-19 pandemic, and although we held on for as long as we could to make the decision to postpone it, we ultimately had no choice but to do so. We had the option of going online in 2020, and then again in 2021 when we tried once more to hold the conference, but we did not want to lose the "family" feeling that has been such a central part of the conferences over the past thirty-seven years since the first conference was held at Exeter University in 1985.

As we entered 2022, there was hope as we saw the light at the end of the tunnel of the pandemic, and as we thought that we would finally be able to hold the face-to-face conference that we had been dreaming of, we were denied this by the global situation of the war in Ukraine and the slowness of the Japanese government to open the borders. With a very heavy heart, we made the decision to hold the conference in an online format, and through using the Gather.Town platform, we believe that we have been able to capture some of the "family" feel that we wanted to maintain, and at the same time open up new possibilities to hold the conference in an innovative and unique way.

Sadly, for many presenters, the work that they had prepared in order to present in 2020 was getting dated, and they were not present at the conference this year. Others did not feel comfortable with the fact that the conference had gone online and withdrew their presentations. Given the uncertainty of the situation with the conference and the fact that we had to suddenly change to an online format—albeit a unique one—we are even more thankful to our many presenters who have put such efforts into preparing for this conference.

This Conference Proceedings is also the hard work of so many people, not the least, the anonymous reviewers who gave such valuable feedback on the submissions, and in particular to Wenzheng Huang, Linh Pham, Bao Nguyen, Andrew Barnes, Yurika Ito, and Cameron Flinn. We would particularly like to thank Cameron Flinn and Yurika Ito for their enormous efforts in copy editing the manuscripts as we prepared for publication.

The International CALL Research Conference is the culmination of planning and hard work by so many people, and the local organising team at Waseda University has been invaluable in making the event successful. More than anything, we would like to thank our presenters and participants for making the conference what it is. Without their innovative ideas and their dedication to research and practice in using technology in language teaching and learning, the conference just would not be possible, so we would like to extend our sincerest gratitude to you all.

> Jozef Colpaert Yijen Wang Glenn Stockwell

Antonie Alm

The University of Otago, New Zealand

antonie.alm@otago.ac.nz

Actualizing the affordances of Machine Translation Tools for language learning

Bio data



Antonie Alm (PhD, UCLA) is an Associate Professor at the University of Otago in New Zealand, where she teaches German language, Computer-assisted language learning and Intercultural Communication. Her current research explores informal engagement with language and the pedagogical use of machine translation in language education.

Abstract

This paper investigates the use of machine translation tools (MTT) for reflective language learning. Drawing on Bowker's (2020) concept of machine translation literacy, it suggests that machine translation tools can support learners in their language development if used critically. It argues that language learners should be supported to seek out affordances of MMT to develop machine translation literacies for specific language skills (reading, writing, listening and speaking) and vocabulary building. This paper presents the preliminary findings of a small-scale study involving ten advanced German language students who explored the features of MTT to meet their individual learning goals for reading. The process of affordance development was analyzed through the lens of Affordance-Actualization Theory.

Conference paper

Introduction

Driven by powerful AI, neural machine translation (NMT) is undoubtedly the smartest kid on the (CALL) block. It has been predicted that machine translation will be able to compete with human translation in the near future (Crossley, 2018). Recent studies indicate that the lexical and grammatical quality of NMT output already outperforms intermediate-level language learners (Lee, 2022). Freely available, online translators provide an impressive range of features, such as Google Translate's conversation mode, enabling people to have bilingual conversations, or its dictionary, which not only translates but also corrects input, as well as provides the option of saving words and phrases for later revision. Further, in addition to document translations, browser extensions translate webpages, making the translation process unnoticeable to the reader. The versatility of machine translation tools (MTT), such as Google Translate, DeepL or Reverso and their omnipresence in digital environments make machine translation an appealing resource for language learners. In fact, recent studies indicate that more than 90% of language students use MTT (Kok Wei, 2021). Survey studies show that students commonly use MTT to look up the meaning of words for reading, translate texts, learn vocabulary, check grammar for writing and pronunciation of unfamiliar words (Briggs, 2018; Kok Wei, 2021; Nugraha et al., 2019). Concerned about the negative impact of MTT use on learning, many teachers ban or restrict their use (Grove & Mundt, 2021) or provide training to ensure that students learn to use them critically and appropriately (O'Neill, 2019).

Machine Translation Literacy

Bowker (2020), a translation scholar, uses the term machine translation literacy to emphasize that machine translation needs to be used skillfully in professional and academic contexts. As she put it, "using machine translation is easy. Using it critically requires thought" (p. 28). Critical use of machine translation requires translators to have an understanding of the functionalities of different online translators and their respective suitability for different tasks. Further, the use of machine translation needs to be transparent and acknowledged. The issue of academic integrity is one of the main concerns voiced by language educators who fear that students may present MT output as their own work (Grove & Mundt, 2021). However, recent studies, including Alm and Watanabe (2021), suggest that most language learners are well aware of the detrimental effect of machine translation on their learning if it is used for unreflective copy and paste translation rather than for reviewing their own texts written in the L2, or making stylistic improvements. Building on this study and the finding that many advanced and experienced learners use machine translation strategically to improve their writing, the current study explores in more detail the conditions under which learners seize the affordances of machine translation for language learning. Drawing on Affordance-Actualization Theory, it seeks to gain an understanding of the processes leading to machine translation literacy for language learning.

Affordance-Actualization Theory

Recently introduced to the area of CALL by Tanaka-Ellis (2022), Affordance-Actualization Theory was originally developed by Strong et al. (2014) in Information Systems to provide a framework for affordance actualization to explain the affordance potential of technologies in organizational contexts. The authors describe affordance as "the potential for behaviors associated with achieving an immediate concrete outcome and arising from the relation between an artifact and a goal-oriented actor or actors" (p. 70). With the term actualization, they refer to "the actions taken by actors as they take advantage of one or more affordances through their use of the technology to achieve immediate concrete outcomes in support of organizational goals" (p. 70).

Tanaka-Ellis (2022) has shown how Affordance-Actualization Theory can be used as an analytical tool in CALL contexts. She points out that unlike earlier views of affordance (e.g. Gibson, 1979; Van Lier, 2004), the model developed by Strong et al. (2014) focuses on the process of actualizing, providing insights into the affordance potential of a technology and its relevance for language learning. The model captures "affordances in transition" (p. 17) by identifying how people engage with a feature of a technology to meet a specific need. Using the terminology of Affordance-Actualization Theory, two elements, 1) the feature of the technology and 2) the characteristics of the actors (users), give rise to an affordance, described as immediate concrete outcomes. In Strong et al.'s (2014) model, these outcomes provide the organization with feedback on the affordance potential of a technology, which may or may not support the organizational goal. The level of actualization is reached when the actions align with goals of the organization. In an exploratory educational context, as presented in this paper, the feedback provided by learners on affordance potential can be used to inform educational goals. To explain, the actions taken by learners might not support language learning (e.g. use of MT to avoid cognitive engagement). However, they might confirm anticipated actions (e.g. use of MT to correct learners' L2 texts) or provide examples of affordance potential for learning that had not previously been identified (e.g. use of audio files to improve pronunciation). An analysis of affordance actualization of individual learners can

therefore help to establish a best practice repertoire to support machine literacy for specific language skills.

This study uses the model of affordance-actualization to analyze how language learners develop machine translation literacy in L2 reading. The larger project also covers L2 writing, listening, speaking, and vocabulary development, examining the interconnections between these affordances. The overarching research question is as follows:

To what extent can self-determined use of machine translation tools support the development of machine translation literacy in advanced language learners?

Method

Participants and task

The participants of this study were ten advanced German students who, after having set themselves learning goals aligned with the five competencies of the Common European Reference Framework, engaged in reading, viewing and listening of self-selected articles, videos, and podcasts from the political youth magazine *Fluter*. They reflected on their learning experiences and shared them with other class members in weekly blogs and Flipgrid videos over a period of six weeks. In their written reflections, students were asked to describe their strategic use of Google Translate and DeepL and their perceived benefits for their language development. In addition, they produced in pairs a repertoire of MTT practices for language learning.

Task design

The design of the learning environment is based on the principles of Self-Determination Theory (Alm, 2006; Ryan & Deci, 2002), supporting learners' basic needs of competence, autonomy and relatedness. The ability to resort to MTTs for learning activities (reading, writing, speaking and listening) was hypothesized to provide confidence, which, with developing machine translation literacy, would support a feeling of competence in learners. The learning environment was designed to be autonomy-supportive, as learners were able to engage with learning materials (text, audio, video) of their choice, made linguistically accessible through MTTs. Learners were asked to share their learning activities with each other to foster a feeling of togetherness.

Data collection and analysis

The learner data (blogs, videos, and written reports) were collected, transcribed (videos) and translated, and then collated on Excel for the preliminary analysis. The data were sorted thematically according to skills and affordance actualization. For the purpose of this report, only the data on L2 reading were used to create a five-step Affordance-Actualization model, providing insights into the process of individual and goal-directed affordance creation for MTT-based reading.

Analysis and Discussion

Table 1 shows the analysis of affordances for machine translation supported L2 reading. The first column lists the features of a machine translation tool that can be used to translate L2 text. The list is not exhaustive, only including the features identified by the participants. They used the Chrome extension of Google Translate, which translates entire webpages, and also glosses the text. By clicking on words, a pop-up menu appears with the written translation and a sound file of the word. The glosses appear on both the original and the translated page. Both Google Translate and DeepL were used as dictionaries on the browser or the app. Finally, the camera on the phone was listed as a feature to take photos of texts, which were then translated by the app. The second column contains the characteristics of the actors, that is, students using machine translation for reading. The list describes their abilities to operate the technical features from column 1. The elements in both lists are needed to *give rise to an affordance*, represented in the third column. Here we find examples of action. Prior to the intervention, participants predominantly used the dictionary function to translate words,

and the camera for live translations of longer texts or to translate posters, as illustrated in these two quotations:

- Last year I was reading a German novel, and I used the camera function to help me. For example, I read it in German first and then held my phone over a page so I could read it in English. After reading it in English, I read it again in German.
- When we were in Germany for the exchange, we both saw some posters. We could use the camera function on the poster to see what it was about.

The fourth columns list the actions taken after participants reflected on their learning goals and on ways to achieve them.

- I translate the Fluter articles into English so I can get an overview of what they're about.
- First, I read the website in German. Then I use Google Translate to read it in English. Finally, I read it again in German.
- The ability to easily translate back and forth means that I can read particularly difficult sentences in English, but still read most of the article in German.
- I write the words I don't know in Google Translate and then also in a Quizlet. This helps me a lot because I see these unknown words twice in German and English, then again when I continue reading the article, so I feel like I actually understood it.

These goals are individual to each learner, fulfilling a concrete purpose. Not all were anticipated as educational goals (column five) but supported the overarching aim of developing machine translation literacy for L2 reading.

Elements giving ris	g rise to an affordance:	nts giving rise to an affordance: 3. Example immediate concrete		4. Goal directed actions needed to actualize an	5. Applicable goals and organizational
1. Features of Machine Translation Tool	2. Characteristics of actors	outcome from data	affordance	context	
Affordance 1 Using	g machine translatior	n for L2 reading	Actualization		
Webpage translation	 know how to install extension, activate translation 		 read in English to get an overview, then read in German read first in German, then in English to fill the gaps read in German, difficult passages in English 	Goals: - MT literacy for L2 reading meaningful engagement with L2 texts - supporting vocabulary development	
Glosses	 know how to activate glosses use sound for pronunciation 		 look up a few words listening and reading for vocabulary learning 		
Dictionary	 know how to enter words by typing, copying, speaking 	 looking up words 	- look up and save words (GT, Quizlet)	_	
Camera	 know how to use camera feature with MT app 	 reading posters reading novel with camera 	 not useful for reading as texts are online 	-	

Table 1. Analysis of affordances for machine translation supported L2 reading

The model clearly shows a shift in MT practices for L2 reading. Prior to the reflective activity, MT was used primarily as a dictionary. In addition, the camera feature was used to translate text. This feature was not perceived as useful in the academic context. Dictionary use, however, was extended to use the feature to save words (creation of phrasebooks) and export the spreadsheet generated by Google Translate into Quizlet for vocabulary revision. Further, the webpage translation and glossing feature (Chrome extension), which had not been used by any learner previously, was used extensively, providing individual learners with support at their level. While some used the translated page to get an overview of the article before they read the original German in detail, others flicked back and forth between versions, or only used the glosses to fill the gaps. The individual approaches show the action potential of the features, which turned into different affordances for each learner. Overall, however, they all contributed to higher engagement with the L2 texts, which would have been too difficult to read for most without the support of the online translators.

This short analysis demonstrates individual pathways to machine translation literacy, which have informed the overarching goal of supporting the development of MT literacy for L2 reading. The news articles became accessible and ensured meaningful engagement with the text. In relation to the SDT framework, the reading activity supported learners with a sense of autonomy, as they were able to choose and understand an article of their personal interest. The online translator further fostered a feeling of competence, as learners felt supported in the task of choosing and reading the text. Finally, the readings were discussed amongst students, giving the reading a social purpose and bringing them closer as a group, thus supporting a sense of togetherness.

Conclusion

This short extract of the preliminary analysis of affordance actualization in MT-based L2 reading has provided some insights into the benefits of MT use in language learning. Students felt supported in the task and, more fundamentally, in their basic human needs of autonomy, competence, and relatedness. The analysis has shown that language learners have a basic understanding of how machine translation can be used for language learning. By aligning MT use to specific learning goals (e.g. ability to read more in the L2), they are encouraged to look for features to support their goal. Extending Bowker's definition of Machine Translation Literacy, it seems appropriate to include the need for technical awareness of MT features which present differently in different MTTs on the browser and on the app, and a sense of curiosity to make them work for their needs. Individual learners have actualized different affordances, providing a pool of best practice examples for machine translation literacy for L2 reading. Being an iterative process, the educational goals will evolve with the developing needs of learning and the technical developments of MTTs. The larger study addresses affordance actualization for writing, listening, speaking and vocabulary development. It is anticipated that the analysis will reveal the interconnections between different affordances, which also support the interconnected nature of the individual tasks in the language learning process.

References

- Alm, A. (2006). CALL for autonomy, competence and relatedness: Motivating language learning environments in Web 2.0. *The JALT CALL Journal*, 2(3), 29-38.
- Alm, A., & Watanabe, Y. (2021). Online translators for L2 writing: a comparison of student and teacher perspectives. CALL and professionalization: short papers from EUROCALL 2021, 23-28.
- Bowker, L. (2020). Machine translation literacy instruction for international business students and business English instructors. *Journal of Business & Finance Librarianship, 25*(1-2), 25-43.
- Briggs, N. (2018). Neural Machine Translation tools in the language learning classroom: Students' use, perceptions, and analyses. *JALT CALL Journal*, *14*(1), 3–24.

- Crossely, S. (2018). Technological disruption in foreign language teaching: The rise of simultaneous machine translation. *Language Teaching*, *51*(4), 541–552. https://doi.org/10.1017/S0261444818000253
- Gibson, J. J. (1979). The theory of affordances. The ecological approach to visual perception. In *The People, Place and, Space Reader* (pp. 56-60). Routledge New York and London.
- Groves, M., & Mundt, K. (2021). A ghostwriter in the machine? Attitudes of academic staff towards machine translation use in internationalized higher education. *Journal of English for Academic Purposes*, 50, 1-11. https://doi.org/10.1016/j.jeap.2021.100957
- Kok Wei, L. (2021). The use of google translate in English language learning: How students view it. International Journal of Advanced Research in Education and Society, 3(1), 47–53.

https://myjms.mohe.gov.my/index.php/ijares/article/view/12459

- Lee, S.-M. (2022). An investigation of machine translation output quality and the influencing factors of source texts. *ReCALL*, *34*(1), 81–94. https://doi.org/10.1017/S0958344021000124
- Nugraha, G., Ratnawati, R., & Surachmat, A. (2019). Exploring low and high students' perception on engaging e-dictionary in mastering vocabulary: Cross-sectional survey. *Indonesian EFL Journal*, 5(1), 37–44. https://doi.org/10.25134/ieflj.v5i1.1609
- O'Neill, E. M. (2019). Training students to use online translators and dictionaries: The impact on second language writing scores. *International Journal of Research Studies in Language Learning*, 8(2), 47–65. https://doi.org/10.5861/ijrsll.2019.4002
- Ryan R. M., & Deci E. L. (2002). An overview of self-determination theory: An organismic dialectic perspective. In E. L. Deci & R. M. Ryan, (Eds.), *Handbook of self-determination research*. Rochester: University of Rochester Press.
- Strong, M.D., Volkoff, O., Johnson, S.A., Pelletier, L.R., Tulu, B., Bar-On, I., Trudel, J., & Garber, L. (2014). A theory of organization-HER affordance actualization. *Journal* of Association for Information Systems, 15(2), 53–85. https://doi.org/10.17705/1jais.00353
- Takana-Ellis, N. (2022). Exploring a Smart CALL environment through the critical lens of Affordance-Actualization Theory. In J. Colpaert & G. Stockwell (Eds.), SMART CALL: Personalization, Contextualization & Socialization (7-28). London: Castledown Publishers.
- van Lier, L. (2004). The ecology and semiotics of language learning: A sociocultural perspective. Norwell, MA: Kluwer Academic Publishers.

Artem Zadorozhnyy, Yu Baohua

The Education University of Hong Kong, Hong Kong

s1130750@s.eduhk.hk, baohuayu@eduhk.hk

Uncovering the role of learning ecology in explaining students' engagement in informal L2 learning activities in digital online environments

Bio data



Artem Zadorozhnyy is a fourth-year PhD candidate at the Education University of Hong Kong. His research interests encompass the development of students' linguistic and digital literacy via integrating innovative multimodal techniques and the role of motivation in predicting positive academic and linguistic outcomes. Artem is a novice researcher who has been previously published in the *ReCALL* Journal and the conference proceedings of EUROCALL 2021 and CALL 2020.



Baohua Yu is an Assistant Professor at the Education University of Hong Kong. Her research interests involve acculturation, learning motivation and strategies in SLA, and digital language learning and teaching. She has published in international refereed journals such as Journal of Multilingual and Multicultural Development, Language Teaching, Higher Education and International Journal of Intercultural Relations.

Abstract

Given the expanding scale of Internet access and ample language learning opportunities associated with the emergence of new digital resources and out-of-class digital environments, SLA researchers started to pay more attention to the language learning experiences of students beyond the classroom level (Lai et al., 2018; Soyoof et al., 2021; Toffoli, 2020). Particular interest is aligned to the perception that while formal education may not satisfy all individual expectations, informal language learning practices might grant students opportunities to engage in cross-cultural communication (Lee & Lee, 2021) and construct their own self-directed authentic learning situations (Lai & Zheng, 2018). Expanding this line of thinking, learning patterns are known to be greatly determined by the members of students' social milieu who can exhibit a favorable or detrimental impact (Niemiec & Ryan, 2009). Ryan and Deci (2016), for instance, numerously stressed out that the nature of relationships between language teachers and students could have an impact on satisfaction of students' fundamental psychological needs and as a consequence influence students' motivation, affecting students' engagement in learning activities. However, researchers underline that EFL students' exposure to authentic language patterns and direct interaction with other people in their online vicinity should be scrutinized more closely nowadays (Noels et al., 2019).

Owing to the complex mechanisms behind these processes and interactions, we aim to apply the sub-theory of Basic Psychological Needs (BPN) to consider the impact of significant others (i.e., language instructors, peers, and target language community members) and the perspective of learning ecology (Barron, 2006) to focus on the

apparent differences in contextual factors pertaining to formal and informal language learning environments. Based on these frameworks, the data obtained via the mixed-methods research approach is set to provide valuable conclusions that could shed some light on the determinants of EFL students' informal language learning activities among prospective EFL teachers of one Central Asian country, Kazakhstan.

Conference paper

Introduction

With the emergence of digital tools that provide students with many possibilities to engage in second language (L2) learning activities, today's research trends in the field of computer-assisted language learning (CALL) are greatly aligned with the experiences of L2 learners in out-of-class language learning environments (De Wilde et al., 2020; Soyoof et al., 2021). Given the current trajectory, in contrast to preceding stages of CALL-related studies that encompassed *structural, communicative,* and *integrative* stages of CALL research, contemporary CALL research agendas focus on L2 acquisition from an ecological standpoint (Chun, 2019). Ecological perspective of human exchanges posits that all habitats where the learning processes transpire can be discussed and analysed in terms of the combination of activities, resources and interactions that occur (Barron, 2006). In other words, an ecological perspective recognises that technologies have enormous educational potential, the influence of which could vary based on students' (Kramsch, 2008; Sefton-Green, 2004).

One of the theoretical frameworks that could be applied along these lines to better comprehend complex mechanisms of human growth in relation to their surroundings is the self-determination sub-theory of Basic Psychological Needs (BPN) (Ryan & Deci, 2017). In particular, BPNs theory will be employed in this paper to differentiate the perceived impact of significant others (i.e., language instructors, peers, and target language community members) on students' engagement in out-of-class language learning activities.

Literature review

8

Self-determination theory (SDT) is a motivational theory at the core of which there is a supposition that every human being has his or her own self-driven propensities to develop (Deci & Ryan, 1985). These propensities could be influenced by various social-contextual factors that either maintain and encourage or interfere with the development of individuals by influencing their level of motivation. The SDT-driven inquiries are centred around several sub-theories which are formed to address various phenomena associated with motivational orientations (Adams et al., 2017). In particular, the theory of basic psychological needs (BPN) considers to what extent social environments and significant others enable individuals to flourish. The BPN theory asserts that the degree of people's basic psychological needs satisfaction is positively connected to the degree of their intrinsic motivation (Agawa & Takeuchi, 2016). Hence, according to the theory, one of the most crucial points for educators is to foster the growth of students' intrinsic motivation because it guides students to explore more, to be engaged in learning activities and to become lifelong learners. Owing to the intrinsically driven nature of informal and out-of-class activities (Toffoli, 2020), the theory of BPNs appears especially applicable to the context of our study.

Drawing on the SDT, the authors assert that in order to stimulate intrinsic motivation of students, educators should establish social contextual conditions that maintain students' satisfaction of three fundamental psychological needs: competence, relatedness, and autonomy (Niemiec & Ryan, 2009). The need for competence refers to desire to be effective in the course of interaction with the social environment while exercising and

employing the gained set of skills and knowledge (Ryan & Deci, 2000). The need for relatedness implies a need to be affiliated to a particular group as well as a desire to be recognized by someone within the group (Deci & Ryan, 2004). At last, the psychological need for being autonomous implies individuals' willingness to be in charge of their own actions and desire to obtain the highest degree of freedom for performing specific actions (Deci & Ryan, 2004). In case when all fundamental needs are fulfilled, an individual has a positive impulse to function due to the augmented level of motivation. On the other hand, if one of the indicated needs is limited, as in the classroom with an authoritarian teacher, students will be lacking the need for being autonomous and, therefore, their motivation to learn will be decreased (Alm, 2007).

While researchers commonly underline satisfaction of BPNs in connection to formal language learning environments and formal language learning processes (Alm, 2006, 2007; Dincer et al., 2019; Noels, 2009; Shelton-Strong, 2020), the nature of online encounters remains less studied. Hence, owing to apparent differences pertaining to formal and informal L2 learning spaces, we decided to conceptualise perceived support of BPNs as a dichotomous construct depicting the impact of significant others (i.e., language instructors, peers, and target language community members) in two kinds of language learning spaces. In particular, while social conditions in formal classroom settings could be considered based on the relationships of students with university/school instructors and peers, digital environments are similarly known to provide students with opportunities to interact with people of various backgrounds and origins.

To some extent, our assumptions corroborate ideas of Noels et al. (2019), who claimed that students' online exposure to native and non-native language speakers should be rigorously scrutinized. Following this line of thought, we aim to expand existing research inquiries to obtain response to our research question: to what extent social contacts in informal and formal environments are conducive to students' engagement in out-of-class language learning activities?

Methodology

For the purpose of answering our research question, we use the data collected during a mixed-methods PhD research project amongst undergraduate students of seven higher education institutions of Kazakhstan. Our study's sample consisted of 472 valid questionnaire responses and 32 online interviews (25 female; 7 male) with students enrolled in a Foreign Languages program aimed at preparing future English language instructors. Due to word limitations, this report will focus entirely on qualitative study findings.

Following the guidelines of Kirsch (1999), all data collection procedures were performed in accordance with the standard ethical considerations set by the Human Research Ethics Committee of the Education University of Hong Kong. As such, all students were given informed consent before participating. Participants were notified about a voluntary participation and their rights to withdraw from the study. Similarly, all required steps were done to protect the participants' wellbeing and ensure confidentiality. For these purposes, all names were substituted with pseudonyms while information about research sites was hidden.

Interviews were set to help students evaluate social conditions in formal university settings to determine the extent to which university teachers support their BPNs and the connection of psychological needs to informal learning experiences. Likewise, similar questions were raised to identify whether encounters in online spaces could be conducive to students' language learning motivation and informal L2 learning activities. To find response to our research question, initial coding procedures were conducted in accordance with the categories underlined by Davis and Bowles (2018), Noels et al. (2019), and Wisniewski et al. (2018). That is, interview coding procedures depicted the

fulfilment of basic psychological needs on a continuum ranging between informative to negative feedback (competence needs), interpersonal immediacy to interpersonal distance (relatedness needs), autonomy support and autonomy control (autonomy needs). In attempts to extend analysis, we made efforts to compare and contrast perceived impact on BPNs across formal and informal online spaces.

Research findings

10

Competence support: from informative to negative feedback

To begin, among all three dimensions of contextual support, most commonly, interview respondents provided information about the perceived support or hindrance of their need for competence. Specifically, the majority of all valid nodes for contextual support fall under the abovementioned category among which the influence of L2 instructors, classmates and L2 community members were noted in relatively similar proportions. The most notable aspects influencing students' satisfaction of fundamental needs were linked to the impact of educators in formal learning settings. It was observed that students whose L2 instructors were reported to place more emphasis on meaning-focused activities exemplified more diverse examples of their engagement in both receptive and interactive informal learning approaches. Along with it, students whose teachers exhibited more attempts to integrate Internet-based and multimedia resources into their teaching methods tended to have higher levels of awareness on how to embed new digital resources and strategies into their daily lives.

Similarly, the role of classmates has received substantial attention, with all students (n = 31) (except one third-year male student) recognising the involvement of their peers in the fulfilment of their competence needs, hence, resulting in their engagement in informal learning practices. Remarkable difference, however, was observed between female and male interviewees in connection to how they perceive language learning support from their peers. Specifically, only male subjects (n = 6) mentioned the presence of unspoken competition between fellow students which was perceived as a motivating factor to actively engage in L2 learning in both formal and informal spaces.

At last, students recognised the essential role of various L2 community members to the satisfaction of their competence needs. To a great extent, interview respondents denoted several groups of influencers, namely: friends, users of communicative-based applications and users of social media communities of interest. It is worth noting that students who reported their frequent engagement in productive informal learning practices repeatedly acknowledged numerous instances of becoming friends with other users who they encountered in different language exchange applications (i.e., Tandem, KakaoTalk, HiNative, HelloTalk). In multiple examples, student teachers revealed that their participation in communicative situations in online spaces provided them opportunities to receive feedback regarding their vocabulary, speaking skills, grammar, and pronunciation.

Relatedness support: from interpersonal immediacy to interpersonal distance

Although the competence support was the most commonly addressed category, research participants extensively shared their views on the support of relatedness and autonomy. As the need for relatedness implies a mutual and reciprocal sense of belonging within a group of people, students' sense of relatedness was addressed based on the perceived behavioural patterns and actions with three groups of social agents, namely, teachers, classmates and the L2 community.

The most frequently mentioned group of social agents was classmates with whom interview subjects were reported to be regularly engaged in various communicative situations and problem-solving tasks both inside and outside classroom environments.

However, respondents acknowledged the role of other social agents as crucial sources of interpersonal immediacy and extensively described situations in which they were offered support and felt the sense of belonging:

We have a very good [university classroom] group, we always support each other. In our group, people really want to learn English and they always bring something new: some words, ideas... When you have a friendly atmosphere, you are not afraid to practice the language. You have a mutual goal with them [with classmates]. (Elena)

During the first months of our studies, [name of the university instructor] was very understanding and adjusted her teaching approach, so that it became much easier for us after some time [to understand the material]. Later it became easier for us to work together. And despite having different levels in the beginning, now we are trying to be on the same page. (Angelina)

I use English for all my social media posts. I know that this is a better approach than using Russian, because people are used to receive explanation of new vocabulary in Russian and always wait for it. But when you post only in English, they will look for answers by themselves and remember them. I think this is much more efficient. Basically, I make different quizzes in English and ask my subscribers to guess the correct meaning of words. When I post a story on Instagram about my life, I always put some English words or hashtags there, because I believe this is useful not only for me, but also for my friends. Well, they always tell me this themselves. (Aigerim)

While watching YouTube videos and [English language] tutorials, there was always a feeling that the authors are interested and actually care that more people can learn English. (Moldir)

Provided experiences of students illustrate how different groups of social agents might impact their sense of connectedness. The first quote exemplifies that a friendly and safe environment in formal classroom spaces and encouragement from peers might facilitate students' interest to work on their language skills. The second quote extends this line of thinking and demonstrates that the role of teacher appears essential in building the community of practice. The quote of a female research informant (Angelina) depicts the importance of teachers' efforts in helping students to form close bonds and forming a small community of L2 learners.

The third excerpt alludes to the significance of perceived interpersonal immediacy by demonstrating that closeness to peers with whom students share a common goal might result in provided competence support in informal L2 environments. Similar instances suggested that encouragement from peers might result in incorporating new language learning experiences into students' toolkit. Lastly, the fourth-year female student (Moldir) provides an illustration of when language learners may experience a sense of relatedness to social agents with whom they do not have direct contact. Specifically, during the interview, Moldir argued that the structure of YouTube tutorials combined with the encouraging and motivating attitude of vloggers and YouTube teachers might contribute to students' sense of proximity.

In contrast to reported cases of interpersonal immediacy, we identified eight valid responses from five students which could be categorised as examples of interpersonal distance. In particular, Mark, a second-year male student, mentioned instances when his need for relatedness has been thwarted by his classmates and L2 university instructors. The quotes of a respondent suggest that he experienced the cases of unwelcome attitude of teachers and the lack of approval from the side of his peers. Interestingly, despite a more negative spectrum of emotions associated with students' unpleasant experiences,

both sorts of cases were found to be conducive to students' engagement in informal learning. This result enables us to draw conclusions about the multifaceted nature of relationships between relatedness needs and participation in informal L2 activities, which should be examined in further research. However, noting students' comments about being related, heard or supported by others, we could conclude on the impact of the sense of relatedness on students' motivation to learn the language with no regard to students' year of study, gender or age.

Autonomy support and autonomy control

Conversely to previous subsections where the roles of different social agents were addressed, students' autonomy support and its control were influenced solely by the university language instructors who were responsible for structuring the formal education process. By and large, interview subjects revealed that autonomy support positively impacts both students' L2 learning motivation and willingness to be engaged in informal language learning activities.

When describing their university experiences, students tended to provide comparisons of their autonomy levels suggesting discrepancies between schools and tertiary institutions environments. That is, the respondents acknowledged that the level of autonomy support at universities was drastically higher in comparison to formal school environments. Students pointed out that successful university language instructors should customize existing assignments and tasks for establishing conditions that favour students' freedom to share personal opinions. By doing so, teachers can give students more control over their learning experience which might result in narrowing the gap between informal and formal language learning practices.

While the most participants were satisfied with the degree of provided autonomy, eight students (25%) were less positive regarding the provision of autonomy support from their language instructors. One student (Sultan) noted that the degree of the autonomy support during formal classroom sessions was somewhat limited, "When we are given a task, there is no complete freedom of action. We are given the usual grammar and vocabulary tasks, and there is not much flexibility in there. There is nothing like that in them, everything is simple and structured." Suggested example exhibits the case when autonomy control might result in limiting students' perceptions about diversification of formal classroom practices. Similar view was echoed by another informant, a third-year female student:

I cannot say that there is a complete autonomy of practice in our university. I thought so before, but teachers seem to be attached to the textbooks. You could count on the fingers of one hand those teachers who really let us speak [freely] and with whom we are not afraid to talk, with whom we are not afraid that we would be judged or given a bad score. (Elena)

Together, these opinions provide important insights into how some students perceive detrimental effects of excessive degree of structure in educational processes. Students' examples demonstrated that in cases when L2 language instructors endorse their choice and freedom of expression, students tend to have augmented levels of intrinsic motivation which could subsequently lead to higher rates of engagement in out-of-class language learning.

Summing up, students mostly agreed that a greater degree of autonomy control could positively influence their language learning motivation and desire to explore "digital wilds". However, given that only six students indirectly suggested the presence of connection between their engagement in informal learning practices and instructors' autonomy support, we could report a weak link between two aspects based on our qualitative dataset.

Conclusions

To sum up, our interviewees appraised the impacts of friends, users of language exchange applications and members of social media communities of interest on their satisfaction of BPNs. To a great extent, the perceived sense of mastery (competence support) of the target language was attributed to the received direct and indirect feedback regarding participants' vocabulary choice, speaking skills, grammar and pronunciation. Interestingly, as opposed to formal university environments, most reported instances associated with receiving feedback during informal learning activities were not connected to students' direct encounters and communicative behaviours. That is, a considerable proportion of IDLE acquirers reported that the amount of content generated by other Internet users and L2 community members provided them with immense opportunities to extract responses to their questions of interest without explicit communication.

The abovementioned point extends the arguments raised by Lai (2017) and Lai and Zheng (2018) who claimed that informal learning allows students to receive immediate assistance from L2 community members. Specifically, our findings support the claim that the availability of authentic language content in online environments allows students to elicit information without initiating communication, but rather through observation and consumption of multimodal content.

Our study confirmed the assumption of the relevance of considerations concerning perceived variations in contextual elements relevant to formal and informal language learning environments. We can conclude that considering not only the influencing roles of teachers and classmates with whom students interact in formal settings but also the roles of L2 community members with whom learners interact through productive (i.e., communication) or receptive (i.e., observation, reading and listening) out-of-class language learning experiences is critical. However, further studies are advised to be conducted for untangling complex relationships between basic psychological needs fulfilment and EFL students' engagement in informal learning practices.

References

- Adams, N., Little, T. D., & Ryan, R. M. (2017). Self-determination theory. In M. L. Wehmeyer, K. A. Shogren, T. D. Little, & S. J. Lopez (Eds.), *Development of self-determination through the life-course* (pp. 47–54). Springer.
- Agawa, T., & Takeuchi, O. (2016). Re-examination of psychological needs and L2 motivation of Japanese EFL learners: An interview study. Asian EFL: Journal Professional Teaching Articles, 89. Retrieved from http://hdl.handle.net/10112/9878
- Alm, A. (2006). CALL for autonomy, competence and relatedness: Motivating language learning environments in Web 2.0. *The JALT CALL Journal*, 2(3), 29–38. Retrieved from https://journal.jaltcall.org/storage/articles/JALTCALL%202-3-29.pdf
- Alm, A. (2007). Motivationstheoretische Grundbedingungen für den erfolgreichen Einsatz von neuen Medien im Fremdsprachenunterricht [Basic conditions of motivational theory for the successful use of new media in foreign language teaching]. *Zeitschrift Für Interkulturellen Fremdsprachenunterricht*, *12*(1). Retrieved from https://ojs.tujournals.ulb.tu-darmstadt.de/index.php/zif/article/viewFile/279/271
- Barron, B. (2006). Interest and self-sustained learning as catalysts of development: A learning ecology perspective. *Human Development*, 49(4), 193–224. https://doi.org/10.1159/000094368

Chun, D. M. (2019). Current and future directions in TELL. *Journal of Educational Technology & Society*, 22(2), 14–25. https://www.jstor.org/stable/10.2307/26819614

- Davis, W. S., & Bowles, F. (2018). Empowerment and intrinsic motivation: A self-determination theory approach to language teaching. CSCTFL Report, 15, 1–19.
- De Wilde, V., Brysbaert, M., & Eyckmans, J. (2020). Learning English through out-of-school exposure. Which levels of language proficiency are attained and which types of input are important? *Bilingualism: Language and Cognition*, 23(1), 171–185. https:// doi.org/10.1017/S1366728918001062
- Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2), 109–134. https://doi.org/10.1016/0092-6566(85)90023-6
- Deci, E. L., & Ryan, R. M. (2004). *Handbook of self-determination research*. University Rochester Press.
- Dincer, A., Yeşilyurt, S., Noels, K. A., & Vargas Lascano, D. I. (2019). Self-determination and classroom engagement of EFL learners: A mixed-methods study of the self-system model of motivational development. SAGE Open, 9(2), 1–15.
- Kirsch, G. (1999). Ethical dilemmas in feminist research: The politics of location, interpretation, and publication. SUNY Press.
- Kramsch, C. (2008). Ecological perspectives on foreign language education. Language Teaching, 41(3), 389–408. https://doi.org/doi:10.1017/S0261444808005065
- Lai, C. (2017). Autonomous language learning with technology: Beyond the classroom. Bloomsbury Publishing.
- Lai, C., Hu, X., & Lyu, B. (2018). Understanding the nature of learners' out-of-class language learning experience with technology. *Computer Assisted Language Learning*, 31(1–2), 114–143. https://doi.org/10.1080/09588221.2017.1391293
- Lai, C., & Zheng, D. (2018). Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL*, *30*(3), 299–318. https://doi:10.1017/S0958344017000258
- Lee, J. S., & Lee, K. (2021). The role of informal digital learning of English and L2 motivational self system in foreign language enjoyment. *British Journal of Educational Technology*, *52*(1), 358–373. https://doi.org/10.1111/BJET.12955
- Niemiec, C. P., & Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education*, 7(2), 133–144. https://doi.org/10.1177/1477878509104318
- Noels, K. A. (2009). The internalisation of language learning into the self and social identity. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 295–313). Multilingual Matters.
- Noels, K. A., Adrian-Taylor, S., Saumure, K., & Katz, J. W. (2019). Motivation and the support of significant others across language learning contexts. *Journal for the Psychology of Language Learning*, 1(1), 106–141. http://www.jpll.org/index.php/journal/article/view/noelsetal
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68.
- Ryan, R. M., & Deci, E. L. (2016). Facilitating and hindering motivation, learning, and well-being in schools: Research and observations from self-determination theory. In K. Wentzel & D. B. Miele (Eds.), *Handbook of motivation at school* (pp. 96–118). Routledge.
- Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Publications.
- Sefton-Green, J. (2004). *Literature review in informal learning with technology outside school*. Futurelab.
- Shelton-Strong, S. J. (2020). Advising in language learning and the support of learners' basic psychological needs: A self-determination theory perspective. *Language Teaching Research*, 1(23), 1–23. https://doi.org/10.1177/1362168820912355
- Soyoof, A., Reynolds, B. L., Vazquez-Calvo, B., & McLay, K. (2021). Informal digital learning of English (IDLE): A scoping review of what has been done and a look

14

towards what is to come. *Computer Assisted Language Learning*, 1–27. https://doi.org/10.1080/09588221.2021.2021242

Toffoli, D. (2020). Informal learning and institution-wide language provision: University language learners in the 21st century. Springer Nature.

Wisniewski, T., White, C., Green, C., Elder, A. F., Sohel, S., Perry, N. E., & Shapka, J. D. (2018). Supporting students through role redefinition: A self-determination theory perspective. *Education as Change*, 22(1), 1–23. http://doi.org/10.25159/1947-9417/3700

Bao Nguyen

Waseda University, Graduate School of International Culture and Communication Studies, Tokyo, Japan

ngdthienbao@fuji.waseda.jp

The role of technology and technology training in language teachers' professional development in the private sector

Bio data

Bao is currently a Ph.D. student at the Graduate School of International Culture and Communication Studies, Waseda University. His research field is applied linguistics, particularly computer assisted language learning (CALL) and teaching English to speakers of other languages (TESOL). His current study is on language teachers' well-being in Vietnam and Japan and the importance of CALL to language teachers' professional development in the two countries.

Abstract

Although the importance of teacher education in CALL has been studied in recent years, how language teachers learn technologies at their workplace and whether the education helps them advance their careers are not yet fully understood. With technology integration becoming a popular teaching approach and the Ministry of Education and Training (MOET) implementing the project "National Foreign Language Project 2008-2020" in Vietnam, it is necessary to study these two subject matters to provide sufficient support on both micro and macro levels, which hopefully may help boost their confidence and enhance their occupational development. Moreover, regardless of the high number of Vietnamese teachers of English in the private sector in Ho Chi Minh city, little is known about how they incorporate technology and its importance to their career pursuit. Therefore, to gain more insight, a case study focusing on the private sector was employed using a mixed-method approach. Data collected from classroom observations and semi-structured interviews were analyzed thematically. A questionnaire survey later was administered to lend more support to the final findings. Final results suggested that although teachers found the provided technologies challenging sometimes, they expressed their enthusiasm toward the learning process. They also regarded their institution not only as a workplace but also as a community where they learn about technology integration from their colleagues and enhance their future career prospects. Lastly, the relationship between teacher education in technology and professional development is also evident on three different levels, namely individual, peer, and institutional levels.

Conference paper

Introduction

Research into language teachers' perspective towards computer-assisted language learning (CALL) is not uncommon. Previous empirical studies have suggested that while integrating CALL into language classrooms may have positive impacts in terms of language teaching and learning (Drossel et al., 2016), it may also create more burden on teachers when sufficient training and support are not provided (Stockwell, 2012). However, even though these studies (Drossel et al., 2016; Stockwell, 2012) are important for understanding language teachers' attitudes toward technology, they were

often carried out in the public sector or public university institutions. It can be seen that language teachers' perspectives in institutions in the private sector is not yet fully understood.

Regarding Vietnamese EFL context, the "National Foreign Language Project 2008-2020 Project" has brought about changes to not only the language teaching field but also to research related to language teachers, especially enquiries regarding difficulties they may encounter from 2008 to 2020. For instance, language barrier (Le, 2007), teaching resources (Le, 2012), and ineffective teaching methodologies (Le & Barnard, 2009) have been evident in previous studies. However, little is known about how to support language teachers and their professional development in both the private and public sectors.

Finally, when it comes to CALL and technology integration, training is usually the commonly debated topic as to which kind of training is sufficient and effective has not yet been agreed on among scholars. In entering their working environment, language teachers play several roles, some of which are time managers, motivators, and technical supporters (Son, 2018). However, these roles are not yet studied thoroughly as determining factors in CALL and language teachers' professional development in their workplace. To support this, Evans et al. (2011) argue that if technology affordances, when enhanced at both individual and social levels, can be helpful to language education, particularly to improving engagement in the workplace.

Therefore, the study was set out to understand how language teachers in the EFL context of Vietnam perceive and execute technology integration in their workplace. Technology training and its importance to teachers' professional development are also another topic of interest as the insight can be beneficial to both teachers and language institutions in the private sector.

Methodology

Using convenience sampling, two teachers were employed for classroom observations and individual post-observation semi-structured interviews. They both recently graduated from universities with a bachelor's degree in English language teaching. In terms of their teaching experience, they had only been teaching for nearly a year by the time the study was conducted. Data gathered were then analyzed thematically and triangulated with results collected from a questionnaire administered to teachers in the same institution. Findings were discussed in terms of three main categories: (1) teachers' perspectives toward CALL; (2) teachers' perspectives toward technology training and learning in their workplace; and (3) teachers' perspectives toward the role of CALL in their professional development.

Results and discussion

Data analysis suggested language teachers' perspectives toward CALL and its role in their professional development can be observed in two interconnected layers, namely individual and social levels of an ecological environment through three relationships, which are teacher-student, teacher-teacher, and teacher-institution relationships. This is in line with the suggestion made by Colpaert and Stockwell (2022) about the socialization of a "Smart CALL" environment where teachers, students, and institutional managers are communicating and influencing one another.

At the individual level, the "SMART CALL" environment helps bring more positive thinking to the teachers and improve the relationship with their students. Regarding their attitudes, the language teachers in this study perceived CALL as a challenge to their career. However, they did not feel intimidated but enthusiastic and eager to learn and apply it to their language classrooms. The obstacles that pertain to CALL and technology were actually motivators, driving them toward self-education and their innovations in teaching approaches. The workplace in this context has been helpful in providing a learning environment for the teachers to learn about CALL and improve the practices of technology integration. Observations have shown that the teachers used technologies quite often during their class time even though the teachers later mentioned in the interviews that they did not have enough experience with technologies. Moreover, according to the interview data, because of this environment, the participants were able to gain more confidence and feel more welcome to the technologies at the institution regardless of their lack of experience in CALL, which eventually would improve their career prospects. The CALL environment also provided teachers more useful affordances to connect, interact, and communicate with their students. Classroom observations have also shown that both teachers used technologies and the provided applications at the institution to interact and establish good rapport with their students in both teacher in that if it had not been for this healthy relationship, they would not have been able to continue their professions.

At the social level, the "SMART CALL" learning environment in this study helped strengthen the bonding among teachers. Both teachers mentioned that peer support was indispensable to their professional development. Through this community, they were able to share their experience in CALL with their colleagues and discuss their anxiety as novice teachers. This helped boost their confidence in language teaching and feel more assertive toward their careers. In other words, the "SMART CALL" environment in this research "reduces the workload and cognitive overload for teachers and learners, reduces their anxiety, increases their self-efficacy, and satisfies their need for autonomy, relatedness, and competence" (Colpaert & Stockwell, 2022, p.3). Moreover, the teacher-institution relationship was also evident in this study as to how the participants discussed their expectations toward the institution in terms of CALL training. Findings obtained from interviews and a questionnaire indicate that teachers expect to have more training on the technologies employed at the institution. However, these expectations had not yet been met. For this reason, this study suggests that communication between teachers and stakeholders, especially institutional management, is of great importance as to fulfill the "SMART CALL" environment, where teachers have more space to convey their needs and institutions have more opportunities to listen to their employees. This, as a result, may help schools in the private sector devise a more reasonable training program accordingly, which contributes greatly to language teachers' professional development.

References

- Colpaert, J., & Stockwell, G. (Eds.) (2022). Smart CALL: Personalization, contextualization, & socialization. London: Castledown Publishers. https://doi.org/10.29140/9781914291012
- Drossel, K., Eickelmann, B., & Gerick, J. (2016). Predictors of teachers' use of ICT in school – The relevance of school characteristics, teachers' attitudes and teacher collaboration. *Education and Information Technologies*, 22, 551–573. https://doi.org/10.1007/s10639-016-9476-y
- Evans, K., Waite, E., & Kersh, N. (2011). Towards a social ecology of adult learning in and through the workplace. In M. Malloch, L. Cairns, & K. Evans (Eds.), *The SAGE handbook of workplace learning* (pp. 356–370). SAGE Publications Ltd. https://dx.doi.org/10.4135/9781446200940.n26
- Le, D. (2012). English as a medium of instruction at a tertiary education system in Vietnam. *The Journal of Asia TEFL*, 9(2), 97–122. Retrieved from http://www.asiatefl.org/main/download_pdf.php?i=99&c=1419300205&fn=9_2_ 05.pdf
- Le, V. C. (2007). A historical review of English language education in Vietnam. In Y. Choi, & B. Spolsky (Eds.), *English education in Asia: History and practice* (pp. 167–180). Asian TEFL.

Le, V., & Barnard, R. (2009). Teaching grammar: A survey of teachers' attitudes in Vietnam. *The Journal of Asia TEFL*, 6(3), 245–273.

Son, J.B. (2018). Teacher development in technology-enhanced language teaching. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-75711-7

Stockwell, G. (Ed.) (2012). Computer-assisted language learning: Diversity in research and practice. Cambridge University Press. https://doi.org/10.1017/CBO9781139060981

Bradley Irwin

Nihon University, College of International Relations, Mishima, Japan

irwin.bradley@nihon-u.ac.jp

Using screencasts to create personalized formative feedback in academic writing courses

Bio data

Bradley Irwin earned his M. Ed. in Second Language Education from OISE at the University of Toronto. At present, he is an Assistant Professor in the Department of International Liberal Arts at Nihon University. His research interests include innovative approaches to feedback, language learner identity and autonomy, CALL, and MALL.

Abstract

This presentation explores the effectiveness of using screencast feedback to improve essay composition in an academic writing course. Participants (N = 20) were asked to complete two 1500-word argumentative essays and revise their essays based on two types of formative feedback provided by their instructor. In one group (N = 12), participants were given written feedback while the other group (N = 8) received screencast feedback for their revisions. The essays were analyzed to compare whether the type of feedback influenced the quality of revision and whether students engaged in self-correction. The results showed that participants who received screencast feedback evaluated it more positively than written feedback, completed a higher percentage of revisions, and engaged in more instances of self-correction.

Conference paper

Researchers have highlighted several beneficial aspects of using screencast feedback in lieu of written feedback. For the purpose of this paper, screencast feedback is defined as a desktop video recording of a student's digitally submitted essay, accompanied by audio narration explaining the feedback points. One appealing aspect of using this approach is the ease with which instructors can improve the quality and depth of detail of their comments. Stannard (2012) notes that teachers can more easily elaborate on their feedback by using screencasts. Similarly, Rahman et al. (2014) state that teachers tend to increase the amount of information they provide when screencasting feedback. This is true of the present study as well. On student compositions of approximately 1500 words, the screencasts were generally six to seven minutes in length which ranged from 750 to 825 words. When providing written feedback on essays of the same length, which included notes or short comments in the margins of the page, feedback amounted to between 150-200 words.

Personalization is another benefit of using screencast feedback. There is an emotive element of hearing someone's voice that is difficult to convey through written words. We can hear someone's joy or exasperation by the way that something is said or through one's tone of voice that allows listeners to infer meaning. Hearing an instructor's voice also seems to deepen the bond between instructors and students. Edwards et al. (2012) found that students in an online learning environment preferred screencast feedback to written feedback and were more likely to develop a sense of community belonging through contextual socialization. Other studies have established that using screencast

feedback helps create a rapport between instructors and students which increases motivation and task engagement (Crews & Wilkinson, 2010; Parton et al., 2010). Henderson and Phillips (2015) found that because students viewed screencast feedback as real, honest, and authentic, this personal and individualized nature of the feedback prompted constructive self-reflection. They also found that students described the feedback as supportive and caring, terms that are not often associated with written feedback.

From a teacher's perspective, the ability to save time and reduce workload is another often cited benefit of using screencast feedback (Ali, 2016; Hynson, 2012; Warnock, 2008). However, as Brereton (2018) points out, the benefit of a reduced workload depends on the feedback context. Rather than saving time, the ability to provide more suggestions and examples of solutions to problematic aspects of students' compositions within the same timeframe is also very appealing.

The purpose of this study was to explore the most fundamental aspect of the feedback process. Is student revision influenced by the type of feedback provided? In order to better understand the relationship between feedback type and revision, two research questions were addressed:

- 1. How do students perceive screencast feedback?
- 2. To what degree are students incorporating the feedback from their teacher when revising their essays?

The participants of the study were 20 second year B2-C1 (CEFR scale) English language learners enrolled in an academic writing course at a Japanese university. The participants were randomly assigned to two groups. 12 students were in a control group which received written feedback while 8 students were assigned to the experimental group which received screencast feedback. While it would have been ideal to separate participants into two groups of equal size, this was not possible because of scheduling conflicts. The academic writing course was conducted weekly for 90 minutes during a 15-week semester. The students submitted two 1500-word essays and received formative feedback on the first draft of their essays (once during week 7 and once during week 14). Students then used the feedback to revise their compositions before submitting a final draft. The drafts were then compared, and revisions analyzed, to determine whether the feedback type influenced the revisions.

Based on surveys conducted examining student perceptions of screencast feedback, it was found that students in the control group were significantly more likely to view the academic writing class as too challenging (M = 4.67, SD = 0.65), while students in the experimental group were more likely to view the course level as appropriate to their level (M = 4.00, SD = 1.07), t(18) = 1.74, p = .04). Students in the control group were also significantly more likely to view the course content as too difficult compared to the experimental group (control group: M = 2.83, SD = 0.58, and experimental group: M =2.25, SD = 0.46, t(18) = 2.39, p = .01). Furthermore, students in the experimental group were significantly more likely to report that they felt a close connection with the teacher because of the feedback style (control group: M = 4.58, SD = 0.67, and experimental group: M = 5.00, SD = 0.00, t(18) = 1.75, p = 0.05) and that they felt more encouragement to revise their essays (control group: M = 4.67, SD = 0.49, and experimental group: M = 5.00, SD = 0.00), t(18) = 1.90, p = 0.04). The experimental group were also more likely to respond that they preferred screencast feedback to traditional written feedback (control group: M = 3.34, SD = 1.07, and experimental group: M = 4.50, SD = 1.07, t(18) = 2.39, p = 0.01).

Regarding the degree to which students incorporated the feedback into the final drafts, it was found that feedback type had no insignificant influence. However, students in the experimental group were more likely to engage in self-corrections. An independent-samples t-test was conducted to compare self-corrections in the control and experimental

conditions. There was a significant difference in self-corrections in the control group (M = 0.58, SD = 0.79, and experimental group: M = 2.88, SD = 1.73), t(18) = 4.04, p = 0.01. Therefore, these results suggest that while feedback type may not influence the likelihood of student feedback uptake, screencast feedback does increase the incidences of self-corrections.

In summary, screencast feedback had a positive impact on how students viewed the difficulty level of the course content, the rapport with the teacher, and the level of encouragement to complete the revisions. They also preferred screencast feedback to traditional written feedback. Furthermore, students who received screencast feedback were significantly more likely to revise sections of their essays that were not specifically mentioned during the feedback process.

To conclude, while it was outside of the scope of this research to determine exactly why screencast feedback encouraged students to revise sections of their essays that were not mentioned in the feedback they were provided with, it is encouraging that this behavior increased. Future research could explore the relationship between screencast feedback and self-revision.

References

- Ali, A.D. (2016) Effectiveness of using screencast feedback on EFL students' writing and perception. *English Language Teaching* 9(8), 106-121.
- Brereton, P. (2018). Improving feedback on academic writing: Combining wikis and screencasts. In P. Clements, A. Krause & P. Bennett (Eds.), *Language teaching in a global age: Shaping the classroom, shaping the world* (pp. 356-362). JALT.
- Crews, T., & Wilkinson, K. (2010). Students' perceived preference for visual and auditory assessment with e-handwritten feedback. *Business Communication Quarterly*, 73(4), 399-412.
- Edwards, K., Dujardin, A.-F., & Williams, N. (2012). Screencast feedback for essays on a distance learning MA in professional communication. *Journal of Academic Writing*, 2(1), 95-126. https://doi.org/10.18552/joaw.v2i1.62
- Henderson, M., & Phillips, M. (2015). Video-based feedback on student assessment: Scarily personal. Australasian Journal of Educational Technology, 31(1), 51-66.
- Hynson, Y. T. A. (2012). An innovative alternative to providing writing feedback on students' essays. *Teaching English with Technology*, *12*(1), 53-57.
- Parton, B. S., Crain-Dorough, M., & Hancock, R. (2010). Using flip camcorders to create video feedback: Is it realistic for professors and beneficial to students? *International Journal of Instructional Technology & Distance Learning*, 7(1), 15-23. https://itdl.org/Journal/Jan_10/article02.htm
- Rahman, S., Salam, A. R., & Yusof, M. (2014). Screencast feedback practice on students' writing [Paper presentation]. Asia-Pacific Social Science Conference, Seoul, Korea.

https://www.researchgate.net/publication/289396637_Screencast_Feedback_Practice_on_Students'_Writing

- Stannard, R. (2012). Simple tools that can revolutionise feedback we provide. In L. Morris & C. Tsolakidis (Eds.), *International conference on information communication technologies in education* (pp. xiv-xix). http://www.icicte.org/Proceedings2012/Papers/Keynote2-Stannard.pdf
- Warnock, S. (2008). Responding to student writing with audio-visual feedback. In T. Carter, M. A. Clayton, A. D. Smith & T. G. Smith (Eds.), Writing and the iGeneration: Composition in the computer-mediated classroom (pp. 201-227). Fountainhead Press.

Cao Thi Xuan Tu, Pham Vu Phi Ho Van Lang University, Vietnam

ho.pvp@vlu.edu.vn

Comparing the effectiveness of online and in-class collaborative writing

Bio data



Cao Thi Xuan Tu works as a lecturer at the Faculty of Foreign Languages, Van Lang University. She has taught English for 8 years, and she has experience in teaching English for a variety of learners ranging from young learners to undergraduate students. She specializes in evaluating and designing teaching materials. She is also a reviewer of the International Journal of TESOL & Education. She is implementing studies on Task-based Language Teaching (TBLT), Computer-mediated Tasks, and Collaborative Writing.



Assoc. Prof. Pham Vu Phi Ho, Ph.D., Assoc. Dean of the Faculty of Foreign Languages, Van Lang University, Vietnam. He used to be a Vice-President of Ba Ria - Vung Tau University, and Vice-President and Dean of Faculty of Foreign Languages of Van Hien University, Vietnam. Pham has been published 56 research local and International articles in both Journals (ISI/Scopus-indexed), and 8 books and course-books, 2 course-books were used for undergrad students at HCMC Open University, VN, and one course-book was used for both the undergraduate and graduate level at Lourdes College, Higher

Education Department, Cagayan de Oro City, Philippines. He has international experience in teaching English at Suranaree University of Technology, Thailand; specifically, he worked as a Professor for Gyeongju University, South Korea, in 2014. He is the Vice President for Administrative Affairs of the AsiaCALL and the managing editor of its Online Journal. He is now the Editor-in-chief of the International Journal of TESOL & Education. He is also an editor for the Asian Journal of Applied Linguistics (Scopus), an editor for the World Journal of English Language (scopus), and a peer reviewer for some international Journals indexed in ISI/Scopus such as Computer Assisted Language Learning, Open Sage, and International Journal of Instruction. His main interests include Academic Writing, peer responses, translation, teaching methodologies, and technology-enhanced learning.

Abstract

Collaborative writing (CW) was found to be beneficial to the students, and it has great impacts on the students' writing outcomes. However, few studies have been investigated to see whether there are any differences in the impacts of writing between online and in-class collaborative writing. The present study analyzed 120 argumentative essays from 60 students, 30 from the virtual class and 30 from the in-class writing. The study only limits to three types of sentences written from each essay, such as simple sentences,

compound sentences, and complex sentences. The study found that there were great effects of the CW on the students' writing quality. However, there was no difference in the effectiveness of CW in both traditional classroom and online platforms. The study claims that either platform of the language teaching, CW helps enhance students' writing skills, and virtual classroom is also a great tool for writing activities.

Conference paper

Introduction

The last two decades have witnessed a huge growth in the application of collaborative learning, especially collaborative writing (CW) in second language teaching. The term collaborative writing has been defined as a type of writing that involves co-authors to participate at all phases of the writing process, sharing responsibility and ownership of the final product (Storch, 2018). According to Storch (2018), the growing interest in collaborative writing is attributed to: (1) the shift in the nature of workplace writing when writing tasks are completed in groups rather than individual; and (2) the emergence of Web 2.0 applications such as blogs, wikis, and Google Docs, which have shifted literacy practices, making the production and transferring of texts easier. To date, most research compared the effect of collaborative writing (pair work, group work) and individual writing on learners' writing performance in either face-to-face interaction classroom (Dobao & Blum, 2013; Pham, 2019; 2021; Storch, 2005) or computer-mediated learning environment (Elola & Oskoz, 2010; Hsu, 2019; Kressler, 2009). The results from these researches consistently indicated that this pedagogical approach has numerous benefits such as enhancing learners' reflective thinking, helping learners to improve grammatical accuracy, fluency, and improving content quality of the texts. Despite the positive effects of collaborative writing on both learning contexts, whether technology-mediated collaborative writing has significance over face-to face collaborative writing is still unknown. Few studies are conducted comparing the effectiveness of collaborative writing in traditional learning context and technology-mediated learning environment. This study aims to fill this gap in literature by comparing writing performance of two classes; one adopted online collaborative writing using MS Teams and the other used collaborative writing in a traditional classroom. The findings of this exploratory study will help clarify the effectiveness of collaborative writing in different learning conditions and provide implications for further practice.

Literature Review

Collaborative writing

24

Collaborative writing has been considered as an effective teaching approach and is widely discussed by many researchers and educators (Pham, 2019; Storch, 2011). According to Storch (2011), collaborative writing refers to an activity in which a pair or a group of participants work together to make a common product. Storch also claimed that all the group members share the ownership of produced text and the peer-review activities solely are not considered collaborative writing. During a collaborative writing activity, students not only brainstorm ideas but also discuss, negotiate to focus on a common goal (Storch, 2019). Therefore, when they work together, they can learn from each other, develop their learning abilities, and achieve better results in their final products (Dobao & Blum, 2013; Heidar, 2016; Pham, 2019).

Collaborative learning is underpinned by constructivism and sociocultural theory of cognitive development. The theory highlights the part of interaction and peer collaboration in second language development. According to Dewey (1938), learning is a

social activity in which learners do things together and interact with each other. Vygotsky's Zone of Proximal Development (1978) explained that higher cognitive functions only appear on the social, intermental plane before the psychological, intramental plane. Novice learners build knowledge in cooperation with more efficient individual experts. Language is the semiotic instrument mediating this process while learning is the gradual internalization of socially built knowledge.

Researchers utilizing sociocultural theory in the research of second language learning claim that learners could have an advantageous effect on each other's development as they could act as both novices and experts (Ohta, 2001; Storch, 2002; Swain & Lapkin, 1998). Due to the difference in both strengths and weakness among two learners, they could provide scaffolded support to each other during cooperation by grouping their different resources and achieve a level of performance which is beyond their individual competence level (Ohta, 2001)

MicrosoftTeams (MS Teams)

MS Teams is an application in the Office 365 ecosystem, which provides users with an effective virtual learning and meeting environment (Tran & Nguyen, 2021). The application was introduced and launched globally on March 14, 2017. After more than six years, MS Teams has gained more than 120 million users worldwide. In the context of Vietnam, MS Teams has become a popular application used for online teaching during Covid-19 pandemic among many schools and educational institutions.

MS Teams provides users with a range of functions supporting learning and teamwork. Teachers and students of a class are assigned to a particular channel in order to have online meetings. Groups of students in the same organization can also create their own channel for teamwork and group discussion. This application allows teachers and students to share materials and documents. It integrates many other Microsoft applications such as Microsoft Words and Excel where students and teachers can collaboratively draft and compose documents. During online lessons, the teacher can assign students into break-out rooms, where they can discuss and do the assignment together.

Simple sentence, compound sentence, and complex sentence

A sentence is a group of words that form one or more clauses to express and communicate a complete thought. There are four basic types of sentences in English including simple, compound, complex, and compound-complex sentences. These kinds of sentences are categorized by the type of clauses used to form them.

There are dependent clauses and independent clauses. A dependent clause is formed with subordinators such as when, if, that, or who. A dependent clause cannot stand alone in a sentence because it cannot express a complete thought. In contrast, an independent clause can stand independently to express a thought with a subject, verb, and often a complement.

A simple sentence is created with one independent clause. A compound sentence is a combination of two or more independent clauses. Coordinators, conjunctive adverbs, or a semicolon can be used to join the clauses in a compound sentence. A complex sentence consists of one independent clause and one or more dependent clauses. In complex sentences, the idea in the main clause (independent clause) is more important than that in the dependent clause. The fourth type of sentence is compound-complex sentence which has at least three clauses—two independent clauses and one dependent clause. In this study, compound—complex sentence and complex are sorted in one group, so three

sentence types involving simple sentence, compound sentence and complex sentence are examined.

Previous studies

26

Collaborative writing, especially collaborative writing on second language learning, has been widely investigated and reported in literature. A range of studies has found positive effects of collaborative writing on learner's performance, especially in terms of accuracy. Storch (2011) recommended that collaborative writing activities, which are lectures of academic writing, would provide a decent learning context for students to improve their quality of academic writing in case of careful training designs. Storch (2002) and Shehadeh (2011) found that collaboration can also help students improve their writing in ideas, organization, lexical resources, and accuracy over individual writing. Other researchers concluded the similar findings that group writing activities provide better writing products among learners (Dobao & Blum, 2013; Watanabe & Swain, 2007).

Dobao and Blum (2013) implemented a study to investigate the students' attitudes and perception to collaborative writing in pairs and small groups. There were 55 Spanish learners divided into two groups, a group worked in pairs and the other worked in a team of four. The study showed that the majority of the participants preferred collaborative writing task, and there was a positive effect of collaborative writing on vocabulary and grammatical accuracy of their written products.

Pham (2021) conducted a study to investigate the effectiveness of collaborative writing on students' writing fluency. The participants were English major students in a university in Vietnam. There were 35 students in the experimental group and 27 in the control group. The two groups were asked to write four different writing essays, one paper for the pre-test, two writing assignments during the training as a normal curriculum, and one paper for the post-test. Unlike the control group, the experiment group was assigned to compose two more essays collaboratively for the pre-test and the post-test. The findings of the study showed that collaborative writing helped improve learners' writing fluency in terms of the number of words in both collaboratively written essays and individually written essays. The study also indicated the positive attitude of students on writing in groups and proposed a useful framework for writing teachers to implement in their classroom. Although the result of the study filled the gap in previous studies (e.g., Ansarimoghaddam et al., 2017; Biria & Jafari, 2013; Storch, 2005; Zabihi & Rezazadeh, 2013), the framework of collaborative writing should be tested in large scale context.

Zabihi and Rezazadeh's (2013) used the Abbreviated Torrance Test for Adults (ATTA) to compare the individual students' writing and pair writing in terms of fluency, accuracy, and complexity. There were 92 university students in Iran participating in the project. The results showed that collaborative writing helped improve accuracy of the written texts compared to individual work. In contrast, there was no effect of collaborative writing on fluency. The limitation of the study is that the students worked in pairs.

Along with the development of the internet and technology, web-based language learning and teaching has become popular and drawn attention among scholars. Especially, online writing or web-based writing tasks using Google Docs, Wiki, or some other platforms has been proven to be beneficial to learners. Talib and Cheung (2017) selected and analysed 15 SSCI journals published from 2006 to 2016. They found that collaborative writing has a positive impact on learners' writing performance in terms of accuracy, critical thinking, and motivation. The study also claimed that technology has enhanced collaborative writing tasks.

Ansarimoghaddam et al. (2017) implemented a study to discover the differences in student's interaction between Wiki and face-to-face when they collaboratively made an

argumentative essay. A whole university class of thirty-two students participated in the research. One group discussed and wrote essays on the wiki platform, while the other groups directly made argumentative essays in the classroom. The research showed that the interaction between group members in collaborative writing came up with social interaction, which motivated students to cooperate and learn from each other. In addition, Wiki interaction made drafting and revising phases more effective to perform while face-to-face interaction was easier for the planning phases.

Kessler et al. (2012) explored the changing nature of collaborative writing affected by Web-based writing contexts. There were 38 Fullbright scholars in a Midwestern university participating in the study. The participants used Google Docs to collaboratively plan and report on a research project. The findings indicated that most students focus more on meaning than form and there were changes made in simple errors in form such as spelling and punctuation. Although fewer students correct their grammar mistakes, the changes they made were generally more accurate than inaccurate. Also, the study found that students were enthusiastic and engaged in working collaboratively.

Valizadeh (2022) examined the effectiveness of collaborative writing on Google Docs on 48 Turkish EFL learners' individual descriptive writings. The participants were divided into two groups with 24 students each. The control group experienced individual writing practice with teacher's corrective feedback whereas the experimental group experienced collaborative writing on Google Docs with corrective feedback from teacher. The findings of the independent samples t-test showed that the collaborative writing group. The researcher suggested that the Google Docs writing environment can help improve learner's individual writing skills. Although this provided the evidence of collaborative writing and individual writing.

Bikiwski and Withatage (2016) conducted a study on the impact of web-based collaboration on individual writing with 59 English L2 learners at a university in the US. Both the experimental group (n= 32) and the control group (n=27) completed four in-class web-based writing tasks. The different treatment was that the experimental group worked collaboratively while the control group engaged in the tasks individually. The findings revealed that students who worked in groups gained higher scores in their individual writings compared to those who completed the web-based tasks individually, although there was evidence of a positive effect of web-based tasks on both groups. Bikiwski and Withatage (2016) also proposed a three Teaching Cycle for Web-Based Collaborative Writing: (1) preparation, (2) collaborative writing, and (3) reflection. In addition, the researchers called for more research on the potential benefits of CALL-based collaboration among L2 writers. In response to this call, our study focuses on investigating the difference between online team writing and in-class team writing. Thus, one research question is raised:

Is there any difference between in-class and online collaborative writing in terms of simple sentence, compound sentence, and complex sentence?

Methodology

Context and participants

The current study took place at the Faculty of Foreign Languages of Van Lang University, Ho Chi Minh City, Vietnam. Forty-three students, ages ranging from 19-20, enrolled in the Writing 4 classes participated in the study. Their English proficiencies were equivalent to B1 of CEFR (The Common European Framework of Reference for Languages). The
participants completed all the prerequisite courses Writing 1, 2, and 3 in their English Major program.

In Writing 1, the students learned sentence structures, describing home, persons, and narrating events. In Writing 2, they learned to compose some paragraph genres, such as logical division of orders, process paragraphs, and opinion paragraphs. In Writing 3, the students learned to write academic paragraphs, such as narrative and descriptive paragraphs and comparison/contrast paragraphs. In Writing 4, they were trained to transfer from paragraphs to essays. During the course, they learned how to make their writing united and coherent. In addition, they learned how to compose three genres of essays, such as comparison/contrast essays, cause/effect essays, and argumentative essays. The writing courses lasted for ten weeks, three hours each week.

Research Design

Text analysis was carried out in this research to measure how different the subjects' written performance was in pre-test and post-test. Specifically, the researchers analyzed word count, sentence types, the use of cohesive devices to evaluate the effect of collaborative writing on students' writing in terms of fluency, complexity, and coherence. The pre-test and post-test essays were analyzed to gain data.

Procedure

The study was carried out following three main stages. In the first stage, the students in two intact classes were formed into small groups. The control class consisted of 22 students, five boys and 17 girls, while the experimental group consisted of 21 students, eight boys and 13 girls. Students were allowed to choose the group to work as long as each group has four to five members. In the second stage, the experimental class was trained on how to create a new document in Microsoft Teams for collaboratively writing and editing an essay. Both the control and experimental class were taught the same lessons. The purpose of this course was to train how to compose comparison and contrast essays, cause and effect essays, and argumentative essays. Besides, it aimed at developing students how to present their ideas in essays logically with unity and coherence. In this stage, students in both classes were instructed how to brainstorm ideas for a particular topic, write the essay and how to give feedback. In the final stage, peer feedback skills were also developed for the students to learn how to evaluate their peers' essays. The students in this course met once a week, three hours for each meeting as the normal curriculum of the university. The main difference between the two classes was that the groups in the control group collaboratively write the essays in class, while those in the experimental group work online using MS Teams.

Data collection

Pre-test and post-test essays were collected via an e-learning site which were provided by the school learning management system. Students in the control group were requested to type their essay in Microsoft Words and submit on their e-learning site. Essays had to be submitted after each lesson. However, only argumentative essays were analyzed for research purpose because learners were requested to write argumentative essays in both pre-test and post-test

Findings and Discussion

There were a total of 120 argumentative essays collected from both control and experimental groups to analyze in this research. A total of 30 pretest papers and 30 post-test papers were collected from each group.

Participants of two groups were requested to write an argumentative essay in the pre-test in 60 minutes. The researchers compared 30 pre-test papers of CG with 30 pre-test papers of the EG using the independent samples t-test of SPSS versus 22 to see if there were any differences between students' number of each sentence type before the study.

	М	SD	t	df	p
Sentences					
Simple sentence					
EG	5.88	2.3	1.81	30	0.07
CG	4.31	2.4			
Compound sentence					
EG	4.5	1.5	1.05	30	0.3
CG	3.94				
complex sentence					
EG	6.94	1.7	.67	30	0.5
CG	6.31				

Table 1 Pre-test of experimental	aroun (FG) vs Pre-test of	control aroun	(CG)
	group (LO,		control group	(UU)

* Independent sample t-test

Table 1 depicts the student's written complexity in terms of number of simple sentences, compound sentence and complex sentence in two groups' pretests. On average, the total number of simple sentences from 30 students in control group (CG) was 4.31 (M=4.31, SD=2.4), while the mean score of the number of simple sentences of the experiment group (EG) was 5.88 (M=5.88, SD=2.3). Students in EG tended to use more simple sentences than those in CG. However, there was no significant difference in the amount of composed simple sentences in two groups according to the result of an independent sample t-test (p=0.07; p>0.05). In terms of compound sentences and complex sentences, the results were similar. The average number of compound sentences in EG was 4.5 (M=4.5, SD=1.5) and that of CG was 3.94 (M=3.94). Complex sentence types of EG and CG had a similar pattern with 6.94 and 6.31 (M=6.94, M=6.31) respectively.

There was no significant difference between the number of these sentence types composed by two groups (p=0.3; p=0.5). That is to say, the complexity involving using different sentence types in the learners' written work before the study was not different.

Is there any difference between in-class and online collaborative writing in terms of simple sentence, compound sentence, and complex sentence?

The purpose of this study was to investigate whether there is any significant difference between collaborative writing in face-to-face and online learning environments. In order to find out the answer to this question, the number of different sentence types in the 30 post-test essays of the CG were compared with those 30 in the EG. The researchers only compare post-test performance of the two groups to investigate if there is any difference in their written work after treatment employed during the study.

	М	SD	t	df	p
Sentences					
Simple sentence					
EG	6.13	1.9	-0.73	30	0.46
CG	6.81	3.1			
Compound sentence					
EG	6.31	1.5	3.41	30	0.02
CG	4.38	1.6			
complex sentence					
EG	8.88	2.3	.60	30	0.55
CG	8.25	3.3			

 Table 2. Post-test of EG vs. CG

* Independent sample t-test

30

As shown in Table 2, the mean score of simple sentences in EG and CG was 6.13 and 6.81 respectively. The p value was 0.46 (p=0.46; p>0.05) means that there was no significant difference in the number of simple sentences composed in both groups. The statistics of complex sentences show a similar result. On average, there were about 8.88 (M=8.88) complex sentences written in EG posttest essays, compared with 8.25 (M=8.25) sentences of CG. It can be seen that there was no difference in the number of complex sentences in two groups (p=0.55). Nevertheless, it is noticeable that there were more compound sentences in EG essays (M=6.31) than in those of CG (M=4.38). The result from the independent sample t-test with t(30)= .60, p=0.02 shows that there was a significant difference between the post-test of EG and CG. In other words, online collaborative writing helps increase the number of compound sentences in learners' individual written work.

Conclusion

This research attempted to investigate the effectiveness of collaborative writing on English learners' writing performance and compared collaborative writing in two different learning contexts, in class and online. The findings indicate that CW facilitates learners' writing fluency in the argumentative essay genre in terms of length increase. Complexity of text has also been found to be enhanced. Specifically, there was an increase in the number of compound and complex sentences in both the control and experimental groups. Furthermore, there seems to be no difference in the effectiveness of CW in both traditional classroom and online platforms, which is MS Teams, in this research.

References

- Ansarimoghaddam, S., Tan, H. B., & Yong, M. F. (2017). Collaboratively composing an argumentative essay: Wiki versus face-to-face interactions. *GEMA Online*® *Journal of Language Studies*, 17(2), 33–53.
- Bikowski, D., & Vithanage, R. (2016). Effects of web-based Collaborative writing on individual L2 writing development. *Language Learning & Technology*, 20(1), 79–99.
- Dobao, A. F., & Blum, A. (2013). Collaborative writing in pairs and small groups: Learners' attitudes and perceptions. *System*, *41*, 365–378.
- Dewey, J. (1938). Experience and education. New York: Macmillan
- Elola, I. (2010). Collaborative writing: Fostering foreign language and writing conventions development. Language Learning & Technology, 14(3), 51–71.
- Heidar, D. M. (2016). ZPD-assisted intervention via web 2.0 and listening comprehension ability. *English for Specific Purposes World*, *17*(4), 1–17.
- Hsu, H.-C. (2019). Wiki-mediated collaboration and its association with L2 writing development: An exploratory study. *Computer Assisted Language Learning*, 32 (8), 1–23. https://doi.org/10.1080/09588221.2018.1542407
- Jafari, N., & Nejad Ansari, D. (2012). The effect of collaboration on Iranian EFL Learners' writing accuracy. *International Education Studies*, *5*(2), 125-131.
- Kessler, G. (2009). Student-initiated attention to form in wiki-based collaborative writing. Language Learning & Technology, 13(1), 79–95.
- Kessler, G., Bikowski, D., & Boggs, J. (2012). Collaborative writing among second language learners in academic web-based projects. *Language Learning & Technology*, 16(1), 91–109
- Pham, V. P. Ho. (2019). The effects of lecturer's model e-comments on graduate students' peer e-comments and writing revision. *Computer Assisted Language Learning.* Advance online publication. https://doi.org/10.1080/09588221.2019.1609521
- Ohta, A.S., (2001). Second language acquisition processes in the classroom: Learning Japanese. Lawrence Erlbaum Associates, Mahwah, NJ.
- Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. Journal of Second Language Writing, 20, 286–305. https://doi.org/10.1016/j.jslw.2011.05.010
- Storch, N. (2001). Comparing ESL learners' attention to grammar on three different collaborative tasks. *RELC Journal*, *32*(2), 104–124.
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, *52*(1), 119–158.
- Storch, N. (2005). Collaborative writing: Product, process and students' reflections. Journal of Second Language Writing, 14, 153–173.
- Storch, N. (2011). Collaborative writing in L2 contexts: Processes, outcomes, and future directions. Annual Review of Applied Linguistics, 31, 275–288.
- Storch, N. (2019). Collaborative writing. Language Teaching, 52(1), 40-59
- Swain, M. (2000). The output hypothesis and beyond. In J. Lantolf (Ed.), Sociocultural theory and second language learning (pp. 97–114). Oxford: Oxford University Press.

- Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal*, 82(3), 320–337.
- Talib, T., & Cheung, Y. L. (2017). Collaborative writing in classroom instruction: A synthesis of recent research. The English Teacher, 46(2), 43–57.
- Valizadeh, M. (2022). Collaborative writing on Google Docs: Effects on EFL learners' descriptive paragraphs. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 6(2), 277-287.
- Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.
- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: Collaborative dialogue between adult ESL learners. *Language Teaching Research*, 11, 121–142.
- Zabihi, R., & Rezazadeh, M. (2013). Creativity and narrative writing in L2 classrooms: Comparing individual and paired task performance. *Bellaterra Journal of Teaching & Learning Language & Literature*, 6(3), 29–46.

Christopher Robert Cooper

Rikkyo University, Tokyo, Japan

cooper@rikkyo.ac.jp

A corpus of short YouTube news videos to inform course design and materials development in an EFL university setting in Japan

Bio data



Chris Cooper currently works as an Adjunct Lecturer in English at the Center for Foreign Language Education and Research at Rikkyo University in Tokyo, Japan. He is also a PhD student at Tokyo University of Foreign Studies. He has been teaching English as a foreign language in Japan since 2010 and his current research interests include corpus linguistics, natural language processing and using YouTube videos for language learning.

Abstract

The aim of the current study was to inform the course design of an elective English news listening course at a private university in Japan. YouTube channels from twelve countries across Asia, Africa, Europe and North America were manually selected by the researcher as a source of in class materials and for course participants to view outside the classroom. To create materials for language-focused learning and assess the vocabulary demands of the videos, transcripts from the channels were extracted using the YouTube Data API and Python, and a corpus of 8,286 video transcripts uploaded in 2021 was randomly sampled to represent the channels. The transcripts were cleaned, and frequency lists were created for adjectives, nouns, and verbs at CEFR B1 level and above. In addition, proper noun and multi-word unit frequency lists were created. An online concordancer was created using the open-source tool ShinyConc, so the learners could investigate the usage of the words in the frequency lists by themselves. A Python script was written to assess the lexical coverage of the videos using the CEFR-J wordlist, and the results suggested that learners may need to be at the CEFR B2 level or above to comfortably comprehend short news YouTube videos. Suggestions for future research are made in the paper, and Python code and supplementary data are available at the author's GitHub page (https://github.com/cooperchris17/yt_short_news).

Conference paper

Introduction

For listening courses, Nation and Yamamoto (2012) suggested adapting Nation's (2007) Four Strands theory in the following way:

- 50% of the time allocated to meaning-focused input (including watching videos, listening to stories or taking part in discussions)
- 25% of the time on fluency-development activities
- 25% on language-focused activities

Proceedings of the XXIst International CALL Research Conference

33

A corpus representing the texts used in a listening course could be a useful tool to evaluate if materials are suitable for meaning-focused input and create materials for the fluency and language-focused strands.

Corpus-based materials

According to Chambers (2010), data-driven learning (DDL) involves learners interacting with concordance lines prepared by the teacher or by using concordance software directly to study the patterns of language and work out how words and phrases are used by themselves. One of the pioneers of this approach, Tim Johns, believed that the computer should be an informant, with the L2 user asking questions, noticing patterns and making their own generalisations (Johns, 1991). He also pointed out that the most important DDL tool is the concordancer. In the intervening years, DDL has become an increasingly popular approach. In Boulton and Cobb's (2017) meta-analysis of DDL research, the data from control/experimental group comparisons (d = 0.95) and pre/posttest designs (d =1.50) indicated that DDL has a large effect on learning. In the EFL context of Japan, Mizumoto and Chujo's (2015) meta-analysis also showed positive effects. The studies were grouped by the gains measured in the original study, and it was shown that studies that used a vocabulary test had a large effect size, papers that tested by category, for example parts of speech or basic grammar, or at the phrase level, including TOEIC type items, had a medium effect size. Studies that used a proficiency test, specifically the TOEIC Bridge, had a small effect size.

To do DDL, a corpus and corpus-tools are needed. According to Anthony (2020), programming is useful to understand the limitations of corpus-based tools, and advantageous if researchers want to create their own corpus and deal with the inevitable noisy data. A further advantage of conducting corpus-based research using a programming language, such as Python or R is the replicability of the research methods. On the subject of quantitative SLA research, Gass et al. (2021) call for materials and data to be made available on repositories for transparency and reproducibility, in the spirit of open science.

Lexical coverage

In reading research, Hu and Nation (2000) estimated that most readers would need to know 98% of the vocabulary in a text to comfortably comprehend it. In listening, Van Zeeland and Schmitt (2013) suggested a lower threshold of 95% in most circumstances, or 90% for some L2 users. As a caveat, they noted that 98% is probably more appropriate if high comprehension is necessary. Research on lexical coverage and the comprehension of videos is still limited, but some research has suggested that 90% coverage may be enough to understand a TV program, such as a documentary, without assistance (Durbahn et al., 2020). If the goal is purely to understand the main points of a video, which also includes visual information, then knowing somewhere between 90-95% of the vocabulary could be an appropriate target.

Another question to consider is what word list to use to assess the vocabulary of texts. There has been some debate recently (e.g., McLean, 2018; Stoeckel et al., 2020; Webb, 2021) in the vocabulary research field about what word counting unit to use: types, lemmas, flemmas, or word families. According to Gablasova and Brezina (2021), there is no real lemma debate and 'the advantage of using lemma as a unit is that it is more precise and requires fewer assumptions about the (morphological and semantic) knowledge on the part of learners than other units' (p. 959). The CEFR-J wordlist (Tono, 2020) is lemma-based and was initially based on a textbook corpus of major textbooks used in compulsory education in China, Korea and Taiwan. Then the wordlists were compared with the English Vocabulary Profile (https://www.englishprofile.org/wordlist), and extra words were added. One benefit of the CEFR-J wordlist over other wordlists that are typically split into 1,000-word bands is the interpretability of the CEFR level, as it can

be matched to the proficiency of learners, who are often put into classes by proficiency level.

Video in language learning

According to Mayer et al. (2014) research, audio lectures were more comprehensible for ESL university learners in the U.S. when they were accompanied by video supporting the meaning. Vanderplank (2016) has argued the case that the use of L2 captions generally makes TV and movies more comprehensible for L2 users than viewing without captions. In addition, Baraowska (2020) showed that both L1 and L2 subtitles reduced cognitive load and increased comprehension when watching a 12-minute TV drama clip. In a longitudinal study, Muñoz et al.'s (2021) results suggested that viewing with L1 subtitles resulted in higher gains in understanding the meaning of vocabulary, but L2 captions led to gains in recalling the written form. Some researchers, such as Majuddin et al. (2021) have taken a more specific approach, with their investigation of multiword expression acquisition showing that two viewings was beneficial compared to one, and viewing with captions resulted in higher gains than no captions.

The advantage of using YouTube videos in the classroom is their flexibility. Not only are they easily accessed on any device, but many videos have the option to choose L1 or L2 captions, transcripts are available, and there is a large selection of short videos, increasing the chance of repeated viewings. In addition a great deal of videos feature L2 users, providing a model of English as a Lingua Franca for viewers.

Research Questions:

- 1. What materials can be produced from a corpus of YouTube videos to promote language focused learning?
- 2. At what CEFR level do learners have sufficient vocabulary knowledge to reach the 90-95% lexical coverage threshold?

Methodology

Corpus design, compilation, and sampling

The corpus was designed to represent short YouTube news videos for EFL learners at the intermediate B1 level and above. It was designed to be used in an elective course at a private university in Tokyo, Japan. The pedagogical goals of the course are to develop the participants' ability to comprehend and discuss domestic and international English language news broadcasts and learn about topical issues.

YouTube channels were manually selected by the author with the following criteria. The channels should include English language news videos, they should be of less than 4 minutes in length, include closed captions and transcripts, and several countries should be represented, with only one channel chosen for each country. When the channels had been selected, metadata was extracted for 500 videos uploaded in 2021 from each channel using the YouTube Data API (2022). Then the youtube-transcript-api (Depoix, 2021) was used to extract auto-generated transcripts from each of the channels. It was necessary to specify auto-generated transcripts to avoid downloading transcripts for videos with no English sound. At this point, channels that had no or few transcripts available were excluded from selection and a list of 12 channels was chosen to collect a larger dataset. In the second round of data collection, the metadata for all short videos uploaded in 2021 was obtained for each channel by making a separate request to the API for each calendar month of the year for each channel. This metadata was used to download a plain text file version of all available transcripts for these videos. The specific YouTube channels and number of transcripts obtained are summarised in Table 1.

Channel	Country	2021 Short Videos	Corpus Sample
ABC News	U.S.A.	1,990	720
Al Jazeera English	Qatar	2,507	720
Arirang News	South Korea	2,842	720
BBC News	U.K.	982	720
CBC News	Canada	667	667
CGTN	China	2,024	720
CNA	Singapore	2,515	720
DW News	Germany	419	419
Nippon TV News 24	Japan	876	720
TVC News	Nigeria	2,855	720
WION	India	2,465	720
i24NEWS English	Israel	1,004	720
	Total	21,146	8286

To more evenly represent each channel and reduce the size of the dataset, a random sample of each channel was selected using the Pandas (McKinney et al., 2010) library in Python. As shown in Table 1, for each channel, 720 videos were randomly selected. This meant that most channels were represented by the same number of videos. As only 667 transcripts were available for CBC News and 419 for DW News, all available transcripts were used for those channels. Due to the sampling, the corpus size was reduced from around 7.2 million tokens to around 2.9 million tokens.

Corpus cleaning

A sample of transcripts were read and reviewed and features that were not representative of speech were removed from the plain text transcripts using Python. The transcripts were all lower case with no sentence punctuation. It was found that the only occurrences of periods after a word were those occurring after numbers, these were deleted from the text files. The following three phrases that indicate sounds other than speech; [Music], [Laughter], [Applause] were deleted. It was decided that transcripts containing less than 50 tokens should be deleted as manual checking of those videos indicated that the majority were largely videos that were mainly visuals with subtitles or were advertisements for the news channels. Some words are automatically censored by YouTube and are displayed as [____] in the transcripts. Censored words only appeared in ten videos, upon manual checking, any words that were mis-transcribed were edited in the text files. Whether the words should be censored is controversial because the audio is not censored, so hard of hearing viewers cannot access the same information. However, all other censored words were not amended in alignment with YouTube's censorship policy.

The final point that was noticed during the cleaning process was a range of different words co-occurring with -19 that clearly represented *covid*-19. As this was an important word in the news in 2021, collocations of 19 and -19 were searched for and a total of 204 patterns were replaced with *covid*-19. Some words were clearly non-words, such as *kovit*-19, in other cases, the word was mis-transcribed as an actual word, such as *coffee* 19. After further manual checking, 69 words were replaced with *covid*, which is also regularly used as a standalone noun without the 19, and 46 other cases were also amended, such as *yukovit* 19 to *new covid*-19.

Frequency list construction

To answer research question one, Part-of-speech(POS)-tagged frequency lists were constructed to allow participants in the course the opportunity to learn vocabulary relevant to the genre of news videos at their level. Due to the issue of tagging proper

nouns caused by the lower-case nature of the transcripts, proper nouns were investigated first. After attempting to tag the texts using the 'off-the-shelf' *NLTK* tagger, which uses the Penn Treebank tagset (https://www.nltk.org/api/nltk.tag.html; Bird et al., 2009), other methods were investigated, using *spaCy* (https://spacy.io/) and the *roberta-large-ner-model*, which is a named-entity recognition model available at Hugging Face (https://huggingface.co/Jean-Baptiste/roberta-large-ner-english). The *roberta-large-ner-model* is described as working well with lower case entities and this seemed to be the case when testing several models from the Hugging Face website (https://huggingface.co/) on a small number of texts.

NLT	K	spaCy		Hugging Face	
word	frequency	word	frequency	word	frequency
xi	141	china	3414	china	2763
xinjiang	129	u.s.	3306	u.s.	2712
south	44	israel	2321	covid-19	2152
october	37	news	2168	israel	1968
joe	32	covid-19	2151	taliban	1223
taiwan	27	president	1927	japan	1131
zealand	26	united	1793	united states	1112
november	23	abc	1682	india	1105
kamala	22	korea	1644	chinese	1077
khan	17	south	1614	afghanistan	981

Table 2. Summary of the most frequent proper nouns using three tagging methods

As can be seen in Table 2, the *NLTK* tagger clearly did not tag proper nouns correctly and there is a difference between the frequencies tagged by *spaCy* and the *roberta-large-ner-model*. This is because *spaCy* tags the individual words in multi-word items, such as *United States*, as individual words unless the model is retrained. However, many entities identified by the *Hugging Face* model did contain several words. To use the most frequent proper noun as an example, the only two words tagged by *spaCy* containing *china* were *china* and *chinatown*. With the Hugging Face model, more than 200 unique entities were identified containing the word *china*. Some of these should probably not be tagged as unique entities, such as *china china* and *china japan*. However, the aim of the proper nouns materials was for learners to compare frequent proper nouns between countries and have a long list of proper nouns that could be used by the instructor to prepare listening activities to identify lesser-known proper nous. Therefore, Hugging Face was used to identify frequent entities per channel, as entities such as *united states* seemed more logical to present to learners than *united* as a standalone proper noun.

Due to the proper noun tagging issue, spaCy was used to prepare frequency lists of adjectives, verbs and nouns. Knowledge of the usage of A1 and A2 level words is crucial, due to the large amount of any text covered by these words. However, for texts that have not been specifically created for language learners, B1 and B2 level words are likely to be essential for comprehension. This being the case and due to the target level of the learners (B1+), any words in the CEFR-J wordlist at the A1 or A2 level were deleted from the frequency lists. Horizontal bar graphs displaying normalised frequencies (per million words) and the percentage of documents containing the word were produced for the top 100 most frequent words at the B1 level and above for each part of speech.

Finally multi-word item lists were produced using *AntGram* (Anthony, 2021), which is software designed to produce lists of frequently occurring words, or *n*-grams. The *n* in *n*-gram represents the number of words, for example, *a lot of* is a 3-gram and *for the first time* is a 4-gram. The following parameters were used in the software:

• 3-grams to 5-grams

- Minimum frequency of 20 and minimum document frequency of 20
- Numbers replaced by #
- No open slots
- Top 1000 sorted by frequency

Multi-word item lists are usually manually edited by multiple researchers with set criteria (e.g. Martinez & Schmitt, 2012). In the present study, multi-word items were deleted by the author from the list if:

- They were used in only one channel (e.g., 'the ABC News')
- N-grams that were actually 2 words (e.g., 'the country's')
- Parts (e.g., 'the same time' and 'at the same' were deleted, 'at the same time' was not)
- all 3-grams that included apostrophes (e.g., 'we're seeing', 'we've seen')
- Any n-gram containing 2 numbers and one word (e.g., the # #)
- N-grams with 'the [noun] and' (e.g., 'the pandemic and', 'the world and')
- Any n-grams that were extracted because there was a reoccurring segment in many videos, for example most ABC News videos ended like this:

hi everyone george stephanopoulos here thanks for checking out the abc news youtube channel if you'd like to get more videos show highlights and watch live event coverage click on the right over here to subscribe to our channel and don't forget to download the abc news app for breaking news alerts thanks for watching

The list was edited by the sole-researcher, meaning there was an element of subjectivity. Also, in contrast to the POS-tagged frequency lists, the multi-word unit list was not filtered by CEFR level.

Online concordancer for data-driven learning

To allow the learners to explore the language in the frequency lists and other words and phrases that they found difficult or interesting, a concordancer was prepared. The open-source tool ShinyConc (http://shinyconc.de/) was chosen because it is fairly easy to set up and allows a corpus to be uploaded, which can be deployed online for free using https://www.shinyapps.io/. After it is deployed, the concordancer can be used in any web browser, including on mobile devices. More sophisticated tools may be available, but one benefit of ShinyConc is that learners do not need to download computer software and upload the corpus by themselves like they would with tools like AntConc (Anthony, 2022) or pay for a subscription as they might need to with tools like Sketch Engine (Kilgarriff et al., 2014; http://www.sketchengine.eu/). In addition, the link to the YouTube video can be added next to the concordance line in ShinyConc, so users can view the video containing the concordance line. However, it is not a hyperlink, meaning that it must be copied and pasted into a new browser window. When the concordancer was deployed online, it was necessary to reduce the size to 5,809 files, due to a file limit for Shiny Apps. This means that learners were exploring a sample of the corpus, and not the full 8,286 files.

Lexical coverage

The decision was made to use the CEFR-J wordlist to assess the vocabulary load of the videos. This was mainly because of the interpretability of the results, as the materials described in the current study were designed for learners at CEFR level B1 and above, this could be used as a benchmark to judge whether the materials were appropriate and how much support should be provided. The New Word Level Checker (NWLC) (Mizumoto, 2021) is an excellent resource for checking the lexical coverage of single texts for a number of wordlists including CEFR-J. However, as the corpus in the current study contained 8,286 texts, a Python script was written using spaCy to assess the vocabulary

demands of all the texts in batches. A list of tuples was produced from the CEFR-J wordlist, including each lemma, its part of speech, and CEFR level. The texts were tagged and lemmatised using spaCy and matched with the list of tuples to assign a CEFR level to each word in each text. Following how texts are profiled in NWLC, proper nouns and numbers were counted as known words first, and the cumulative lexical coverage was calculated at each CEFR level, to give an indication of the lexical difficulty of the texts, lemmas that were not in any of the lists were added to an *others* list. Extra lemmas were added to the list of tuples to match how spaCy tags some words. This was done by checking how words were tagged by New Word Level Checker and in consultation with the English Profile (https://www.englishprofile.org/). A full list of the additions with justifications, along with a Jupyter notebook containing the code to calculate the lexical coverage for multiple texts is available on an online repository (see below). The performance of the script written for this project seems to be close to NWLC. However, one area that needs improvement is how proper nouns are tagged. NWLC seems to be more accurate at tagging proper nouns.

Additional materials and data

The main steps of the methodology have been described here. However, more detailed information including Python code, Jupyter Notebooks, and additional results are available on the following GitHub repository: https://github.com/cooperchris17/ yt_short_news.

Results and discussion

Frequency Lists

An example of the part of speech frequency list graphs, and multi-word units list created as classroom materials can be seen in Figure 1 and Table 3. Some of the words in the example in Figure 1 would be classed as A1 words in a different part-of-speech. For example, the noun *test* is an A1 level word, but the verb *test* is a B2 level word. The aim of the lists is for learners to search for words or phrases they are interested in using the concordancer and notice their own patterns of how the word is used in context. As the corpus is genre specific, this may help them understand the language in other news videos that they watch.



Figure 1. Example of part-of-speech wordlist

Original Rank	Multi-word unit	Frequency	Document frequency
2	a lot of	1589	1077
3	more than #	1551	1207
4	one of the	1458	1204
9	going to be	820	620
10	the end of	771	669
11	as well as	733	634
12	# percent of	725	576
14	be able to	708	601
177	this is the	654	586
178	some of the	652	561
179	# year old	643	510
181	in the country	618	519
202	the number of	581	463
203	the first time	567	510
209	re going to	536	433

Table 3. Top 15 frequent multi-word units (not filtered by CEFR level)

For proper nouns, it might be noticed by learners that most countries use proper nouns related to their country or neighbouring countries—for example, *Canada, Ontario* and *Vancouver* in the Canadian news channel shown in Figure 2. Many of the lists also included the major economic countries, the U.S. and China. Most of the frequent proper nouns would probably be known by learners, but some may not be, such as *RCMP* (Royal Canadian Mounted Police). If a learner was interested in news from a particular country, it could be worth becoming familiar with proper nouns such as this.



Figure 2. Example of proper noun frequency list

40

Online concordancer

V T I OI IN 0001

A screenshot of the concordancer created using ShinyConc can be seen in Figure 3. The concordancer (https://coopersensei.shinyapps.io/yt news shinyconc/) is very functional and its greatest benefit is that users can get started with it immediately on any device without the need to download software or upload a corpus. The concordance lines can be ordered by the word to the right of the node word, however, this is not the case for words to the left of the node word. If the word 'left' is clicked, the concordance lines are ordered by the first word in that column, which limits pattern searching to a certain extent. The display of metadata and YouTube links is also very useful. However, as the purpose is to improve listening ability, it would be better if a tool was designed to link to an embedded YouTube video at the point of the concordance line. This could probably be done using JSON files extracted by the *youtube-transcript-api*, which include timestamps for each line. At least two online tools are already available that search YouTube transcripts and display embedded videos. However, either the register of the videos is unclear (https://youglish.com/) or they are fixed to the genre of TED Talks (https://yohasebe.com/tcse/; Hasebe, 2015). A tool loaded with a range of specified genres, or a tool where users could upload their own transcripts would be useful additions to what is already available.

channel mont	h	Display			342 tokens found (2025908 v	ords in select	ion)
ABC News (US) 1 Al Jazeera (QA) 2 Arirang News (KR) 3 BBC News (UK) 4		KWIC O Data Show 10 - entries url 0	month 0	channel 0	left 0	Search center 0	right
Edit Restrictions		youtube.com/watch?v=y6aE9FIFfXk	8	Al Jazeera (QA)	ill nets were put in place to	protect	a rare porpoise that lives on
Search / Filter		youtube.com/watch?v#lzZncnSMfs8	9	BBC News (UK)	en has responded by vowing to	protect	abortion rights he said the I
	0	youtube.com/watch?v=E998vfW6qP0	1	Arirang News (KR)	isasters and a revised law to	protect	abused children the latter re
house	~	youtube.com/watch?v=HE53pliecvQ	7	CBC News (CA)	higher level of antibodies to	protect	against and to fight against
Search mode		youtube.com/watch?v=r197mwzT0Tk	4	CBC News (CA)	en a single dose can not only	protect	against getting covid-19 but
⊖ string		youtube.com/watch?v=ki0QM1/TZIo	2	CNA (SG)	ke rail lines and assets that	protect	against rising sea levels is
word		youtube.com/watch?v=YqbJXN0qSso	12	ABC News (US)	the vaccine and booster shot	protect	against serious illness and d
 regular expression 		youtube.com/watch?v=87AyPbcj_wl	2	CBC News (CA)	azeneca vaccine probably will	protect	against severe disease those
case-sensitive		youtube.com/watch?v=H1hpX8uiRks	10	BBC News (UK)	ct-based journalism served to	protect	against the abuse of power an
		youtube.com/watch?v=ALIRs4ajsdM	2	Arirang News (KR)	s a proportionate response to	protect	american and american interes

Figure 3. Screenshot of ShinyConc

Lexical Coverage

The lexical coverage of the videos including proper nouns and numbers at the CEFR A2, B1 and B2 levels are shown in Figures 6, 7 and 8. The line in the centre of each boxplot represents the median lexical coverage score for the videos in the specified channel. The boxes represent 50% of the data points between the first and third quartiles, the dots to the left of each boxplot show all of the data points, so the dispersion of the results can be clearly seen. At the A2 level, the boxes containing the data points 25% above and 25% below the median fall between around 75% and 85% lexical coverage. At the B1 level these figures rise to just below 85% and around 92%, and at the B2 level between 89% and 95%. Based on these results, it seems that L2 users' level should be at least B2 before they can freely and comfortably watch short news YouTube videos. However, as news videos are often supported by visuals that are likely to aid comprehension and often have the option to display captions, they could still be suitable for B1 level L2 users.

Lexical Coverage at A2 Level

42



Figure 6. Boxplots showing lexical coverage at A2 level



Figure 7. Boxplots showing lexical coverage at B1 level



Figure 8. Boxplots showing lexical coverage at B2 level

If we compare the lexical coverage of the results in this study with the CEFR guidelines related to watching TV shown in Table 4, we can see that the CEFR descriptors suggest that learners can deal with news to a certain extent from the A2 level. It could be the case that setting reasonable learner expectations when listening and teaching listening strategies (e.g., Vandergrift et al., 2006) is equally as important as checking vocabulary. In addition, many other factors have been shown to affect listening comprehension, such

as redundancy, schema, concreteness and orality (Bloomfield et al., 2010). For example, the channel with seemingly the lowest lexical coverage scores, Nippon TV, may be more comprehensible to L2 users from Japan, because they may be familiar with much of the content.

CEFR Level	Can-do descriptor
A2	Can identify the main point of TV news items reporting events, accidents, etc. where the visuals support the commentary.
B1	Can understand a large part of many TV programmes on topics of personal interest such as interviews, short lectures and news reports when the delivery is relatively slow and clear.
B2	Can understand most TV news and current affairs programmes.
Council of E	urope (2020, p. 53)

Table 4.	TV news-related	CEFR can-do	descriptors
----------	-----------------	-------------	-------------

Whether proper nouns and numbers should be counted as known words is also a point of discussion for listening texts. In written texts, proper nouns are capitalised, which may be easier for L2 users to recognise during reading. This is not the case with listening texts and classifying proper nouns as known words has been called into question by some researchers. Kobeleva (2012), who compared groups listening to texts with known and unknown proper nouns, concluded that proper nouns. This seems to be true for many of the videos in the current study (see Figure 9). The teaching of strategies for recognising proper nouns in listening texts could be an activity worth spending time on in the classroom and a potential topic for future research in this area.



Figure 9. Percentage of proper nouns in the corpus by channel

A final point worth mentioning related to lexical coverage is the words that were not tagged as proper nouns, numbers, or A1-B2 words. A total of 5,912 lemma types were in this category, so clearly it would not be advisable to draw attention to many of these words in a classroom situation. However, some of the frequent words that may be considered specialized vocabulary based on current trends and usage could be introduced to learners, such as Covid-19 related words (e.g., *variant*, *vaccinate*, *vaccination* and *booster*), along with register-specific words, such as interjections (e.g., *uh*, *well*, *um*, and *like*). Identifying irregularly high frequency vocabulary is a benefit of reviewing the lexical

coverage of a corpus of texts that are specifically related to the target genre of class content, as it is not possible to assign a CEFR level to all vocabulary, especially when the high frequency may be related to a specific time-period or genre.

Conclusion and future suggestions

It has been shown in the current study that it is possible to extract transcripts from YouTube that are highly related to classroom content and use the transcripts to create materials that can be used in the classroom. There were various obstacles when dealing with the YouTube data, such as the tagging of proper nouns and transcript cleaning. In addition, a script was written in Python to calculate the lexical coverage of a large number of texts using the lemma-based CEFR-J wordlist. For future projects, the lexical coverage calculation could be improved, specifically the way that proper nouns are tagged, or the inclusion of C1 and C2 level words to give a more accurate lexical profile. Triangulating the results with user ratings would also be beneficial to investigate how accurate the CEFR-J wordlist is at predicting the difficulty level of videos. It could be that L2 users at a lower CEFR level than the lexical coverage score could also comprehend news videos due to the visual support. In addition, a concordancer could be developed linking concordance lines directly to their timestamp in YouTube videos. These are points that the researcher would like to investigate in the future, and any interested parties are welcome to view the code and datasets on GitHub (https://github.com/ cooperchris17/yt_short_news), and make comments or get in contact with any suggestions.

References

44

- Anthony, L. (2020). Programming for corpus linguistics. In M. Pacquot & S. T. Gries (Eds.), *A practical handbook of corpus linguistics* (pp. 181-207). Springer.
- Anthony, L. (2021). AntGram (Version 1.3.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software
- Anthony, L. (2022). AntConc (Version 4.0.6) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software
- Baranowska, K. (2020). Learning most with least effort: subtitles and cognitive load. *ELT Journal*, 74(2), 105-115. https://doi.org/10.1093/elt/ccz060
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python: Analyzing text with the natural language toolkit*. O'Reilly Media, Inc. Retrieved from https://www.nltk.org/book/

Bloomfield, A., Wayland, S. C., Rhoades, E., Blodgett, A., Linck, J., & Ross, S. (2010). What makes listening difficult? Factors affecting second language listening comprehension. University of Maryland. Retrieved from https://apps.dtic.mil/sti/pdfs/ADA550176.pdf

- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. Language learning, 67(2), 348-393. https://doi.org/10.1111/lang.12224
- Chambers, A. (2010). What is data-driven learning? In A. O'Keefe, & M. McCarthy (Eds.) The routledge handbook of corpus linguistics. (pp.345-358). Routledge.
- Council of Europe. (2020). Common European framework of reference for ;anguages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing. Retrieved from www.coe.int/lang-cefr
- Depoix (2021). youtube-transcript-api 0.4.2. Retrieved from https://pypi.org/project/youtube-transcript-api/

Durbahn, M., Rodgers, M., & Peters, E. (2020). The relationship between vocabulary and viewing comprehension. System, 88(102166), 1-13. https://doi.org/10.1016/j.system.2019.102166

- Gablasova, D., & Brezina, V. (2021). Words that matter in L2 research and pedagogy: A corpus-linguistics perspective. *Studies in Second Language Acquisition*, 43(5), 958-961. https://doi.org/10.1017/S027226312100070X
- Gass, S., Loewen, S., & Plonsky, L. (2021). Coming of age: The past, present, and future of quantitative SLA research. *Language Teaching*, 54(2), 245-258. https://doi.org/10.1017/S0261444819000430
- Hasebe, Y. (2015). Design and implementation of an online corpus of presentation transcripts of TED talks. *Procedia-Social and Behavioral Sciences, 198*, 174-182. https://doi.org/10.1016/j.sbspro.2015.07.434
- Hu, M. and P. Nation. 2000. Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403–430. Retrieved from https://scholarspace.manoa.hawaii.edu/bitstream/10125/66973/13_1_10125_6 6973_rfl131hsuehchao.pdf
- Johns, T. (1991). Should you be persuaded: Two samples of data-driven learning materials. *English Language Research Journal*, *4*, 1-16. Retrieved from https://lexically.net/wordsmith/corpus_linguistics_links/Tim%20Johns%20and% 20DDL.pdf
- Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, 1(1), 7-36. Retrieved from

https://www.sketchengine.eu/wp-content/uploads/The_Sketch_Engine_2014.pdf

- Kobeleva, P. P. (2012). Second language listening and unfamiliar proper names: Comprehension barrier?. *RELC journal*, 43(1), 83-98. https://doi.org/10.1177%2F0033688212440637
- Majuddin, E., Siyanova-Chanturia, A., & Boers, F. (2021). Incidental acquisition of multiword expressions through audiovisual materials: The role of repetition and typographic enhancement. *Studies in Second Language Acquisition*, 43(5), 985-1008. https://doi.org/10.1017/S0272263121000036
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, 33(3), 299-320. https://doi.org/10.1093/applin/ams010
- Mayer, R. E., Lee, H., & Peebles, A. (2014). Multimedia learning in a second language: A cognitive load perspective. *Applied Cognitive Psychology*, 28(5), 653-660. https://doi.org/10.1002/acp.3050

McKinney, W. (2010). Data structures for statistical computing in python. Proceedings of the 9th Python in Science Conference, 445, 56–61. Retrieved from http://conference.scipy.org/proceedings/scipy2010/pdfs/mckinney.pdf

- McLean, S. (2018). Evidence for the adoption of the flemma as an appropriate word counting unit. *Applied Linguistics*, 39(6), 823–845. https://doi.org/10.1093/applin/amw050
- Mizumoto, A. (2021). New word level checker [Web application]. https://nwlc.pythonanywhere.com/
- Mizumoto, A., & Chujo, K. (2015). A meta-analysis of data-driven learning approach in the Japanese EFL classroom. *English Corpus Studies 22*, 1-18.
- Muñoz, C., Pujadas, G., & Pattemore, A. (2021). Audio-visual input for learning L2 vocabulary and grammatical constructions. *Second Language Research*. Advance online publication. https://doi.org/10.1177%2F02676583211015797
- Nation, P. (2007). The four strands. *International Journal of Innovation in Language Learning and Teaching*, 1(1), 2-13. https://doi.org/10.2167/illt039.0
- Nation, P., & Yamamoto, A. (2012). Applying the four strands. International Journal of Innovation in English Language Teaching and Research, 1(2), 167-181. Retrieved from https://openaccess.wgtn.ac.nz/articles/journal_contribution/Applying_the_four_s
- trands/12552020/1/files/23373002.pdf Stoeckel, T., Ishii, T., & Bennett, P. (2020). Is the lemma more appropriate than the flemma as a word counting unit? *Applied Linguistics* 41(4), 601-606
 - flemma as a word counting unit?. *Applied Linguistics, 41*(4), 601-606. https://doi.org/10.1093/applin/amy059

- Tono, Y. (2020). The CEFR-J wordlist version 1.6. Compiled by Yukio Tono, Tokyo University of Foreign Studies. Retrieved from http://www.cefr-j.org/download.html
- van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension?. *Applied Linguistics*, *34*(4), 457-479. https://doi.org/10.1093/applin/ams074
- Vandergrift, L., Goh, C. C., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language learning*, 56(3), 431-462. https://doi.org/10.1111/j.1467-9922.2006.00373.x
- Vanderplank, R. (2016). Captioned media in foreign language learning and teaching: Subtitles for the deaf and hard-of-hearing as tools for language learning. Palgrave Macmillan.
- Webb, S. (2021). The lemma dilemma: How should words be operationalized in research and pedagogy? *Studies in Second Language Acquisition, 43*(5), 941-949. https://doi.org/10.1017/S0272263121000784
- YouTube Data API (2022) YouTube data API. Retrieved from https://developers.google.com/youtube/v3

Guoyuhui Huang

The University of Hong Kong, Hong Kong SAR, China

hgyuhui@connect.hku.hk

Repetition supports the effects of Involvement Load Hypothesis on improving students' productive vocabulary performance

Bio data



Guoyuhui Huang is a Ph.D candidate at the Faculty of Education, the University of Hong Kong. Her research interests include understanding how technology can be utilized to enhance students' language learning.

Abstract

Productive vocabulary refers to retrieving and applying the words in speaking and writing. It forms the basis for EFL (English as a Foreign Language) learners to express themselves accurately and fluently. Recent years have witnessed a growth of publications examining the effects of the Involvement Load Hypothesis (ILH) on productive vocabulary performance, although with somewhat mixed results. The present study explored whether 'repetition' could complement the ILH in improving EFL learning of productive vocabulary. Correspondingly, two WeChat Applets (Applet 1.0 and Applet 2.0) were designed to help Chinese EFL learners apply productive vocabularies in the IELTS (International English Language Testing System) examination. Applet 1.0 was chiefly developed based on the ILH. Applet 2.0 was developed based on ILH with the additional support of repetitive reading activities. Specifically, learners studied with Applet 1.0 merely encountered each target item once. By contrast, learners studied with Applet 2.0 would meet each required vocabulary eight times. Fifty Chinese college students, divided into a control group (CG, n = 26) and an experimental group (EG, n = 24), participated in the present study. Both groups were asked to write a composition every week while learning with different applets (CG learned with 1.0 and EG studied with 2.0). Three paragraph writing tests, namely pre-test, post-test and delayed-test were administered to assess their productive vocabulary proficiency. We found the EG significantly outperformed the CG in terms of the post-test and delayed-test. Therefore, it was concluded that repetition and ILH were indeed compatible, which could result in better productive vocabulary acquisition.

Conference paper

Introduction

One explicit and well-recognized taxonomy for vocabulary knowledge is receptive and productive knowledge. Receptive vocabulary is often defined as the ability to comprehend the words in listening and reading; while, productive vocabulary requires producing the relevant words in speaking and writing (Schmitt, 2010). Put another way, receptive vocabulary knowledge refers to the ability of a learner to recognize a word, while

productive vocabulary refers to the ability of a learner to apply the word correctly. Since the application of a word is far more challenging than recognition, previous studies (e.g., Malmström et al., 2018) have reported that learning productive vocabulary is more demanding for L2 learners. However, there has been little research on effective instructional strategies to improve productive vocabulary knowledge among L2 learners, Therefore, a critical question remains: how could we help students effectively acquire productive vocabulary?

Literature Review

Previous Studies on Involvement Load Hypothesis

Over the past two decades, a wealth of research (e.g., Bao, 2015; Keating, 2008; Zou, 2017) has empirically showed that productive vocabulary acquisition might hinge on the degree of involvement in processing new words, which Laufer and Hulstijn (2001) called the Involvement Load Hypothesis (ILH). The ILH consists of three essential constructs: need (N), search (S) and evaluation (E). Furthermore, Laufer and Hulstijn (2001) defined some degrees of prominence for each component. Need is a motivational dimension of the hypothesis, which is defined as the learners' intention to understand the words. It is hypothesized to be absent (N-, 0 points), moderate (N+, 1 point) or strong (N++, 2 points). For example, need is absent when a target word is not required to accomplish a task. It is induced to a moderate degree when learning the target words is externally imposed by teachers or other authorities. By contrast, need is strong when the intention to learn the words is self-imposed. Search is a cognitive element. It refers to the attempt to ascertain the L2 form of a word or its equivalent L1 meaning. According to Laufer and Hulstihn (2001), it is absent (S-, 0 points) when no such effort is required (i.e., a reading comprehension task with some marginal glosses), and it is present (S+, 1 point) when students must seek the L2 form or L1 meaning to complete an assignment. With respect to evaluation (a cognitive element), it entails the comparison between a target word and other related words, or knowledge of a word, with the context of utilization to decide to check if it fits. Evaluation appears to be absent (E-, 0 points), moderate (E+, 1 point) or strong (E++, 2 points). It is absent when learners do not need to determine which word or tense of the word to employ. Moderate evaluation requires one to distinguish between the multiple meanings or forms to find the most appropriate one when the context is given (i.e., a gap-filing activity). By contrast, evaluation is strong if learners need to combine the new word with others to create an original context (as opposed to a given). For example, students need to use the target words in writing sentences or compositions.

The sum of the components with their degree is called the task involvement load (IL). Laufer and Hulstijn (2001) claimed a task with a higher IL was more effective and would yield better vocabulary acquisition and greater vocabulary retention. Recent years have observed a flurry of publications examining its predictive power on productive vocabulary acquisition, although with somewhat mixed success (e.g., Alavinia & Rahimi, 2019; Jafari et al., 2018; Tahmasbi & Farvardin, 2017; Zou, 2017). Admittedly, an emerging body of scholars (e.g., Pourakbari & Biria, 2015; Tahmasbi & Farvardin, 2017) still observed a significant decrease of productive vocabulary knowledge in the delayed-test, even when participants conducted the tasks with high IL (i.e., the sentence-writing task or the paragraph-writing task). Therefore, we posited there is room to improve the ILH. In other words, other factors might complement the ILH in improving students' delayed productive vocabulary performance.

Background Literature on the 'Repetition'

48

A recent meta-analysis on ILH (Yanagisawa & Webb, 2021) which scrutinized 42 empirical studies has aligned with our prediction that other factors might support the effects of the ILH. To begin with, they reported that the predictive ability of ILH was limited: it explained merely 15.0% and 5.1% of the variance in effect sizes on the

post-test and delayed-test of vocabulary knowledge, respectively. Moreover, they demonstrated that 'repetition' might positively contribute to immediate post-test learning gains, regardless of the amount of IL of the vocabulary learning tasks. Repetition refers to the number of times a learner encounters an unknown word in contexts. According to Folse (2006), repetition provided retrieval opportunities for students. Put differently, when a student encountered a word repeatedly while reading, he was likely to focus greater attention on that unfamiliar word in the first several encounters to infer its meaning and retrieve information learned about that word from the previous encounters (Pellicer-Sánchez et al., 2021).

However, the repetition effect was not found on delayed post-tests possibly due to the limited number of repetitions in the studies (Yanagisawa & Webb, 2021). The target words in a majority of studies were not repeated; when repetition was included, the mean frequency was quite low (M = 3.69, Mdn = 4) (Yanagisawa & Webb, 2021). If future studies were to improve students' delayed vocabulary performances, a higher number of repetitions might be required.

Contributions of the present study

Given the gaps in the previous literature, an attempt to expand and advance the theoretical and practical utility of ILH in productive vocabulary contexts is needed. An empirical study that overcomes the deficiencies in previous studies (e.g., inadequate repetitions) is also needed.

In the present study, we posited ILH with higher number of repetitions (more than four times, as indicated by the preceding literature review) might result in better productive vocabulary acquisition. To testify the hypothesis above, we designed two WeChat Applets: Applet 1.0 and Applet 2.0. Applet 1.0 was based on the ILH solely. Applet 2.0 was also informed by ILH supplemented with repetitive reading activities. Learners studied with Applet 1.0 morely encountered each target item once, while learners studied with Applet 2.0 would meet each vocabulary eight times. Furthermore, two groups of students: a control group (CG, n = 26) and an experimental group (EG, n = 24) studied with different Applets (CG studied with Applet 1.0 while EG learned with Applet 2.0) to reveal the discrepancies.

Method

Participants

A randomized control trial was embedded within a quantitative research design. A total of 50 EFL students (CG, n = 26; EG, n = 24), whose age range was between 18 and 20, participated in the present study. The Human Research Ethics Committee at the authors' university has endorsed an ethical approval.

Instruments and Materials

WeChat Applet (Applet) is a web application that can be accessed without downloading or installing. It could bring about a rapid transfer of digital data between instructors and learners, and thus has been proved useful for improving students' learning performances (e.g., Wu et al., 2018). Therefore, we used Applets to distribute learning materials in the present study.

Target words

Since the participants in the present study were going to take the IELTS (International English Language Testing System), we selected 40 words at IELTS level. Moreover, we selected ten verbs, ten adjectives, ten nouns and ten conjunctions to control the possible

confusion of part of speech with tasks. Students needed to learn ten target words per week. Accordingly, they finished learning by the end of the fourth week.

Paragraph-writing task

Previous studies have posited that paragraph-writing task was superior to other vocabulary tasks since it involved deeper processing (e.g., Kim, 2011; Zou, 2017). According to Zou's (2017) analysis, when students were writing paragraphs, semantic contexts of target words were generated, which was conducive to word learning. Therefore, every week, we asked students to finish a semantically acceptable and grammatically correct paragraph with ten target words they learned that week. In the end, they needed to submit four compositions in total.

Learning materials in the Applet 1.0

Applet 1.0 distributed the reading tasks for the CG. Since students in the CG encountered each target word once, forty sentences were prepared for them. These sentences were adapted or selected from Oxford Advanced Learner's English-Chinese Dictionary (Hornby, 2009), Merriam-Webster Dictionary (online version, see at https://www.merriam-webster.com) and Cambridge Dictionary (online version, see at https://dictionary.cambridge.org), in which target words were embedded. The teacher-researchers and other five EFL teachers checked the appropriateness of vocabulary and syntax of the sentences for the participants. Figure 1 illustrated the learning interface in Applet 1.0. Each sentence appeared with one target word embedded. After each sentence, the corresponding glosses (parts of speeches and Chinese translations) for that target word were provided in the brackets. We take the word adhere as an example. In this example (see Fig.1), the target word is adhere and its gloss provides its part of speech (a verb), and Chinese translation (粘附, 附着).

Learning materials in the Applet 2.0

50

Applet 2.0 delivered the reading materials for EG. As noted earlier, Applet 2.0 was supplemented with repetitive reading activities. According to Uchihara et al. (2019), although repetition might contribute to incidental vocabulary acquisition, there were diminishing learning gains as the number of repetition increased beyond a certain point (around 20 encounters). This has aligned with other studies claiming that more was better might not always be true (Elgort et al., 2018; Pelliver-Sánchez, 2016). Elgort et al. (2018)'s eye-tracking research has suggested repetition remained a plateau effect on the processing of target words after a certain number of encounters (eight to ten times). Considering all these useful findings, we let students in the EG meet eight times of each target word. Correspondingly, we designed the learning process as follows: each word appeared twice from every Monday to Thursday (2 times * 4 days = 8 times); every Friday was left for students to finish a paragraph writing task.

Therefore, altogether 320 sentences (40 target words * 8 repetitions) were selected from the dictionaries we mentioned above (Oxford Advanced Learner's English-Chinese Dictionary, Merriam-Webster Dictionary and Cambridge Dictionary). They were divided into four sets, with each presented per week. Each set included 80 sentences as well as ten target words. Every day from Monday to Thursday, students met 20 sentences per day, with each target word appearing twice. After each sentence, their corresponding glosses (parts of speeches and Chinese translations) for the target words were provided (see Figure 2).



Figure 1. Learning interface of the Applet 0.0



Measures

We conducted three paragraph writing tests (pre-, post- and delayed-tests) with two demands on students to gauge their productive vocabulary proficiency. First, they were required to write down the Chinese translations of ten pre-selected words. Second, they should write a paragraph using these words as many as possible. One point would be given to a sentence precisely applied to a target word with an accurate Chinese translation. One point would also be assigned if other parts of sentences were wrong (Tahmasbi & Farvardin, 2017). For example, if the target word is *adhere*, students could be given one point if they wrote 'We should adhere to the regulations, no matter what is happened'. This is because the clause which contains the target word is both semantically acceptable and grammatically correct. The maximum score for each pre-, post-, and delayed-test was ten.

Procedure

To begin with, a pre-test was administered before the treatment to see if there was a significant difference of prior knowledge between groups. Subsequently, a four-week treatment was implemented. During the treatment, CG and EG students used Applet 1.0 and Applet 2.0 to study 40 target words, respectively. Moreover, both groups were required to submit a paragraph-writing assignment every week. After four-week learning, students were given a post-test. Two weeks later, a delayed test was administered.

Results

As demonstrated in Table 1, the prior knowledge before the treatment was similar (EG: M = 0.17, SD = 0.82; CG = 0.19, SD = 0.49), p = .89. After the treatment (see Table 2), both groups enhanced significantly in terms of productive vocabulary in the post-test (EG = 7.79, SD = 1.82; CG = 4.69, SD = 2.57) and the delayed-test (EG = 7.79, SD = 1.61; CG = 4.19, SD = 1.70), F (2, 96) = 273.63, p = .00, η^2 = .85). Furthermore, there was a significant main effects of group, F (1,48) = 47.10, p = .00, η^2 = .50. Table 2 also illustrated a significant interaction between time and group, F (2, 96) = 22.44, p = .00, η^2 = .32, indicating the EG significantly outperformed the CG in terms of improvement in productive vocabulary proficiency. The results also suggested an almost large effect size for time and almost medium effect size for group. Furthermore, the results of pairwise

comparison of time and group (see Table 3) suggested a significant improvement of both groups from the pre-test to the post-test, and a sustained knowledge retention of two groups from the post-test to the delayed-test.

 Table 1. Means and SDs for productive vocabulary scores for the EG and CG in the pre-test, post-test and delayed-test

	Pre	-test	Post	-test	Delaye	d-test
Group	М	SD	М	SD	М	SD
EG (n = 24)	0.17	0.82	7.79*	1.82	7.79*	1.61
CG (n = 26)	0.19	0.49	4.69	2.57	4.19	1.70

Note. M = mean; SD = standard deviation

Maximum score = 10

*p < .001

Table 2. Results of mixed ANOVA on productive vocabulary tests

Source	Type III SS	df	MS	F	Significanc	Partial η^2
		Betwe	en-subject			
Group	185.25	1	185.25	47.10	.00	.50
Error	188.78	48	3.93			
Within-subject						
Time	1174.81	2	587.41	273.63	.00	.85
Time × Group	96.33	2	48.17	22.44	.00	.32
Error (time)	206.08	96	2.15			
Nista CC average						

Note. SS = sum of square; MS = mean square

Table 3. Results of pairwise comparison of time and group

Group	(I)time	(J)time MD SE		Significance	
Pre-test		Post-test	-7.63	.45	.00
		Delayed-test	-7.63	.34	.00
EG	Post-test	Pre-test	7.63	.45	.00
		Delayed-test	00	.46	1.00
	Delayed-test	Pre-test	7.63	.34	.00
		Post-test	.00	.46	1.00
	Pre-test	Post-test	-4.50	.43	.00
		Delayed-test	-4.00	.33	.00
CG	Post-test	Pre-test	4.50	.43	.00
		Delayed-test	.50	.45	.27
	Delayed-test	Pre-test	4.00	.33	.00
		Post-test	50	.45	.27

Note. MD = mean difference; SE = standard error

Discussion

The results witnessed a significant performance of EG, since the average score of EG was approximately twice as large as CG in the post-test and delayed-test. The better performance of EG is largely due to the higher number of repetition. Studies (e.g., Webb, 2007) which examined the effectiveness of repetition on productive vocabulary acquisition illustrated that a higher number of repetitions might help students acquire productive knowledge of orthography, grammatical function, syntax and association, which was crucial for accurate language use. Take grammatical function as an example.

Students need to know the word *beauty* is a noun while the word *beautiful* is an adjective. Therefore, they could avoid ungrammatical utterances, such as 'She is a beauty lady.' Moreover, our quantitative results are aligned with the rationale of ILH, since both groups improved significantly better after the treatment. Of particular note, we thought participants in CG might forget how to apply the target words in the delayed-test. Perhaps most surprisingly, their productive vocabulary knowledge is still retained later. It seemed that tasks with high IL could also contribute to long-term retention. In the present study, we were not sure which (repetition or IL) carried more relative weight in terms of productive vocabulary's long-term retention (since EG students also performed well in the delayed-test). Therefore, further research comparing the effects of repetition and IL on vocabulary retention within a single study is required.

Limitations and Implications for Future Studies

First, the present study focused solely on improving students' learning performance. Extant studies have found learning productive vocabulary often frustrated students (e.g., Qian & Sun, 2019). And this may assuage negative feelings or reluctance to learn productive vocabulary in the curriculum. However, the present study did not consider 'motivation' or 'attitude' as variables. More nuance and clarity in how different types of motivational elements advance the effects of ILH and therefore result in better productive vocabulary performance are needed. Second, in a study that extensively relied on visual input, learning gains usually occur through repetitive visual input. Eye-tracking data could shed light on the role of repetition in the productive vocabulary learning process, yet this was not the purpose of the present study. In this respect, future studies should consider analyzing these electronic data to generate more robust and reliable results.

Conclusions

The present study is an attempt to investigate whether a higher number of repetition might bolster the effects of ILH and result in students' better productive vocabulary performance. As can be seen from the results, students appeared to benefit more from the combination of repetition and ILH. This study therefore provides additional evidence that other factors—as opposed to need, search, and evaluation—should be equal (Laufer & Hulstijn, 2001). More research with other proficiency groups is warranted to further refine our understanding of the potential factors contributing to the effects of ILH.

References

- Alavinia, P., & Rahimi, H. (2019). Task types effects and task involvement load on vocabulary learning of EFL learners. *International Journal of Instruction*, 12(1), 1501-1516.
- Bao, G. (2015). Task type effects on English as a foreign language learners' acquisition of receptive and productive vocabulary knowledge. *System*, *53*, 84-95.
- Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, 40(2), 341-366.
- Folse, S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, 40(2), 273-293.
- Hornby, A. S. (2019). Oxford advanced learner's English-Chinese dictionary (7th ed.). Oxford: Oxford University Press.
- Jafari, S., Izadpanah, S., & Rahmani, R. (2018). The effect of task-induced Involvement Load on unfamiliar L2 vocabulary learning: Sentence writing, summary writing, imaginary story writing and creative sentence writing. *Applied Research on* English Language, 7(1), 67-88.

- Keating, G. D. (2008). Task effectiveness and word learning in a second language: The involvement load hypothesis on trial. *Language Teaching Research*, 12(3), 365-386.
- Kim, Y. (2011). The role of task-induced involvement and learner proficiency in L2 vocabulary acquisition. *Language Learning*, *61*, 100-140.
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26.
- Malmström, H., Pecorari, D., & Shaw, P. (2018). Words for what? Contrasting university students' receptive and productive academic vocabulary needs. *English for Specific Purposes (New York, N.Y.)*, 50, 28-39.
- Pellicer-Sánchez, A. (2016). Incidental L2 vocabulary acquisition from and while reading: An eye-tracking study. *Studies in Second Language Acquisition*, *38*(1), 97-130.
- Pellicer-Sánchez, A., Conklin, K., & Vilkaité-Lozdiené, L. (2021). The effect of pre-reading instruction on vocabulary learning: An investigation of L1 and L2 readers' eye movements. *Language Learning*, 71(1), 162-203.
- Pourakbari, A., & Biria, R. (2015). Efficacy of task-induced involvement in incidental lexical development of Iranian senior EFL students. *English Language Teaching*, 8(5), 122-131.
- Qian, Y., & Sun, Y. (2019). Autonomous learning of productive vocabulary in the EFL context: An action research approach. *Digital Scholarship in the Humanities*, 34(1), 159-173.
- Schmitt, N.(2010). Researching vocabulary: A vocabulary research manual. New York: Palgrave Macmillan.
- Tahmasbi, M., & Farvardin, M. (2017). Probing the effects of task types on EFL learners' receptive and productive vocabulary knowledge: The case of Involvement Load Hypothesis. *SAGE Open*, *7*(3), 1-10.
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language learning*, 69(3), 559-599.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.
- Yanagisawa, A., & Webb, S. (2021). To what extent does the Involvement Load Hypothesis predict incidental L2 vocabulary learning? A meta-analysis. *Language Learning*, 71(2), 487-536.
- Zou, D. (2017). Vocabulary acquisition through cloze exercises, sentence-writing and composition-writing: Extending the evaluation component of the involvement load hypothesis. *Language Teaching Research*, 21(1), 54-75.

Hanae Ikeshita

Department of Information and Media Studies, Faculty of Arts and Sciences, Sagami Women's University, Kanagawa, Japan

ikeshita_hanae@isc.sagami-wu.ac.jp

A pilot study of children with dyslexia and learning foreign languages using 3D letters

Bio data

Hanae Ikeshita is an associate professor at Sagami Women's University. She received her Ph.D. in Global Information and Telecommunication Studies from Waseda University in 2011. She has developed literacy-learning methods for children with developmental dyslexia using 3D technology. Her recent research interests include accessible digital media and bilingual digital books to help children with dyslexia learn foreign languages.

Abstract

Children with developmental dyslexia find it challenging to learn a foreign language, as they often also have difficulties reading and writing their native language. In this study, we examined and assessed the intuitive recognition of the structures of Roman alphabetical letter expressions in children with developmental dyslexia to support their English learning. In the assessment, 18 participants who displayed difficulties in reading or writing English were examined in a comparative study of six types of 3D alphabetical letter expressions. The results suggest that the background color and the letter color combined in letter expressions affect the ease of recognizing the stroke orders of alphabets. Additionally, a questionnaire-survey was conducted among children with dyslexia to understand difficulties faced while learning native and foreign languages at school. Ten participants' parents were asked about their children's difficulties when learning native and foreign languages at school. Based on the relationship between Japanese and English acquisition, the survey showed that children with difficulties in reading and writing Japanese Kanji in elementary and junior high school also found it difficult to acquire English. Our results suggest that children who have difficulty acquiring Kanji, especially writing, also face difficulty learning English. Identifying children who experience difficulties in learning Kanji may allow them to receive English-learning support at an early age.

Conference paper

Introduction

Dyslexia is a neurological disorder characterized by difficulties in acquiring reading and writing skills, regardless of a child's age, intelligence, or motivation (Ferrer et al., 2010; Lyon et al., 2003). The comorbidity of attention-deficit hyperactivity disorder (ADHD) and dyslexia is among the common developmental disorders during childhood (Gilger et al., 1992; Kronenberger & Dunn, 2003). In approximately 4% of children with a learning disability (LD), dyslexia and ADHD coexist (Pastor & Reuben, 2008).

A new curriculum was officially implemented in Japanese public elementary schools in April 2020; foreign language (English) became a mandatory subject for students in the fifth and sixth grades. Previously, English was categorized as a foreign language activity,

and its classes included only speaking and listening, but in the new curriculum, reading and writing skills have also been included. Additionally, a foreign language is not easy to learn for children with dyslexia, who also have difficulties reading and writing their native language. Special education classes are conducted in Japanese public elementary schools, but children with dyslexia are often enrolled in regular classes. Hence, they receive education in the same learning environment as other children.

The Japanese writing system consists of three scripts: Hiragana, Katakana, and Kanji. In Japanese public elementary schools, students learn to read and write 46 Hiragana and Katakana characters and 1,006 Kanji characters. A case study of Japanese bilingual children reported that they could read Japanese accurately and fluently but faced difficulties in English reading and spelling (Wydell & Butterworth, 1999). Thus, typically developing children might face similar difficulties learning English as children with dyslexia.

Individuals with dyslexia are likely to have superior levels of various abilities despite problems with the acquisition of reading and writing skills. For example, enhanced visual-spatial recognition has been associated with dyslexia (von Károlyi et al., 2003). Visual-spatial recognition is required to understand the three-dimensional (3D) position of objects; 3D objects contain more information than two-dimensional (2D) objects. Foreign language teaching for children with dyslexia often employs a multi-sensory approach involving the visual and auditory senses. For example, multimedia learning incorporating tablet computers is being widely used in Japanese public schools. Multimedia learning also consists of assistive technology with audio support for children with dyslexia. Knoop-van Campen et al. (2020) suggested that adding audio support compensates for reading difficulties but eventually hampers the learning of students with dyslexia.

In Hiragana, each character represents one of 46 unique syllables. Yamazoe [Ikeshita] et al. (2009) developed a literacy-learning system to illustrate Hiragana using a stereoscopic 3D display; results suggested the possibility of using stereoscopic vision to help individuals with developmental dyslexia write Hiragana characters. However, some studies have found that participants show symptoms such as asthenopia or headaches after viewing a 3D display (Kooi & Toet, 2004; Lambooij et al., 2009; Read & Bohr, 2014); further, stereoscopic displays are difficult to use and thus are rarely used in schools. Primary schools in Japan have recently introduced tablets in class, but little is known about adding 3D visual support in multimedia learning of foreign languages for Japanese children with dyslexia.

In the present study, we explored how this visual support affects children with dyslexia in their ability to recognize stroke order. We attempted to resolve these challenges by using a tablet computer that does not require special equipment to test the efficacy of 3D English learning as a tool in foreign language education. We examined the advantages of 3D alphabetical letter expressions in a comparative study of six types of these expressions to support English education for children with dyslexia. Additionally, we conducted a questionnaire-survey among children with dyslexia to understand the difficulties they face while learning native and foreign languages at school.

Methods

Participants

Participants comprised 18 students with dyslexia and other LDs who had difficulties reading or writing English (aged 12 to 20 years; seven cases with ADHD, 11 cases with autism spectrum disorders). Written informed consent was obtained from adult participants and children's parents. This study was approved by the Ethical Review Committee at Sagami Women's University (No. 19052).

Materials

We used a 3D modeling software (The Foundry MODO 10.2v1, Autodesk Maya 2016) to model 3D Roman alphabetical letters (hereafter "3D letters"), which were exported in COLLADA format for display on tablets. Each 3D letter was divided into strokes positioned at regular intervals in depth according to the stroke order. The 3D letters were developed so that learners could operate them by touching the screen with a finger and rotating them 360 degrees (Figure 1). The 3D letters used Helvetica (Gothic) for the font. To move the 3D letter, the user had to tap the letter with a fingertip; the 3D letter could be moved up, down, left, or right in a 360-degree arc. Each letter stroke was displayed in the order they were to be drawn in.



Figure 1. Example of a 3D letter

3D letter expressions were set to the color values of white (RGB: 255, 255, 255), black (RGB: 0, 0, 0), gradation 1 (black to white), and gradation 2 (white to black); white and black were used for background colors. The color settings for two gradations were applied from the beginning of the first stroke to the end of the last stroke. Six types of 3D letter expressions were created (Table 1).

Туре	Letter color	Background color
А	White	Black
В	Black	White
С	Gradation 1 (RGB value 255-0)	White
D	Gradation 2 (RGB value 0-255)	White
Е	Gradation 1 (RGB value 255-0)	Black
F	Gradation 2 (RGB value 0-255)	Black

 Table 1. Six types of 3D letter expressions

Setting

The experiment was conducted in a private room with a table on which the tablet was placed and a chair for the participant to sit on. The stimuli were displayed on a 9.7-inch tablet (Apple iPad Air) at a 50 cm distance. Participants' behavior during the experiment was recorded using a digital camera. Stimuli constituted the six types of 3D letter

expressions, as shown in Table 1. The letters ``b'' and ``d'' were selected, as children with dyslexia often confuse them.

Procedure

First, participants were informed of the experimental procedure and asked to practice the operation of 3D letters to be presented on the screen beforehand. The two 3D letters were presented simultaneously on the tablet screen, and participants were instructed to move each letter from side to side and up and down using a finger. After watching and comparing the 3D letters, participants were asked, "Which letter's writing order is understandable?"; they selected a letter by pointing a finger. This was regarded as one trial, and a set of 15 trials were performed. The presentation order of the 3D letter types was randomized across participants, and it varied for each of the 15 trials. This experiment examined how to easily recognize the correct stroke order. In addition, the participants were interviewed and asked to comment on the 3D letters.

Second, ten participants' parents were asked four questions (Table 2) regarding difficulties their children faced in acquiring native and foreign languages in school.

Table 2. Question items about the process of children acquiring native and	foreign
languages in schools	

Theme	Question item
Process of acquiring children's native language (Japanese)	 Do you think your child could read and write Japanese (Hiragana, Katakana, and Kanji) in elementary school? Do you think your child experienced distress while learning Japanese in elementary school?
Process of acquiring a foreign language (English)	3) Do you think your child could read and write English in elementary or junior high school?4) Do you think that your child experienced distress while learning English in elementary or junior high school?

Data analysis

To compare the 3D letters, participants' perceptions of readability were analyzed using Thurstone's law of comparative judgment (Thurstone, 1927). Thurstone's method provides one-dimensional data based on preference judgments for two items and can be used to transform rank-ordered data. The most popular simplification is Thurstone's Case V model, which was used in this study because children were able to select stimuli easily. The Case V model assumes all response options have equal variance and zero correlations. The results of the interval scores were plotted on an interval scale.

In the questionnaire-survey, the response data were analyzed to determine the correlation between participants' acquisition of native and foreign languages. Correlations with acquiring different languages were analyzed with Fisher's exact test and Cramer's V.

Results

Results of understanding stroke order

Figures 2 and 3 show the rank-order results for 3D letter expressions. The horizontal axis shows the understanding of stroke order; levels of understanding increase from right to left. The ranking order for understanding stroke order of 3D letters was type A, B, D, E, F; C for the letter "b"; and type A, B, E, D, F and C for the letter "d."



Figure 2. Results of understanding stroke order for "b"



Figure 3. Results of understanding stroke order for "d"

We used the binomial test, to check for significance in the comparison between types of 3D letter expressions. Understanding the stroke order of 3D letters was significant for both letters "b" and "d." Analysis of "b" revealed significant differences between types A and C, and B and C (p < 0.01);as well as A and F, B and E, and C and E (p < 0.05). Regarding "d," significant differences were noted between A and C, and D and F (p < 0.05).

In the interviews regarding 3D letters, participants mentioned, "I think it is easier to learn the alphabet using the 3D letter," "I will be able to memorize the letter more easily with the 3D letter," "It is easy to understand how to rotate the letter," and "I think that the 3D letter makes it enjoyable to learn English."

Results of correlations with learning native and foreign languages in school

Figure 4 presents survey responses regarding acquisition of reading. For Japanese (Hiragana, Katakana, and Kanji), participants sometimes misread Kanji but were able to read Katakana and Hiragana. Ten and nine participants could read Katakana and Hiragana, respectively. Six participants were unable to read English.

Figure 5 presents survey responses regarding acquisition of writing. For Japanese (Hiragana, Katakana, and Kanji), most participants were able to write in Katakana and Hiragana. Nine and eight participants could write in Katakana and Hiragana, respectively. Seven participants were unable to write in Kanji and English.

59

We used Fisher's exact test and Cramer's V as nonparametric statistical methods to test for correlation between difficulty in acquiring native and foreign languages. Cramer's V varies between 0 and 1; a value greater than 0.25 is considered a very strong relationship. Both Kanji and English were strongly associated with writing difficulties, as confirmed by strong Cramer's V correlations (1.00, p < 0.01). English writing and reading were strongly associated with writing difficulties, as confirmed by correlations (0.80, p < 0.05). Writing Kanji and reading English were strongly associated with writing Kanji or reading English difficulties, as confirmed by strong Cramer's V correlations (0.80, p < 0.05).

Results of correlations with distress in learning native and foreign languages in school

Five participants selected "strongly agree" or "agree" regarding experiencing distress while learning Japanese and English in school (Table 3). The mean evaluation (SD) given for learning Japanese and English in school were high, 3.10 (1.37) and 3.50 (1.08), respectively.

We used the chi-square test and Cramer's V to test for correlation between mental stress in learning native and foreign languages. There was no significant difference between learning Japanese and learning English in school (χ^2 = 17.50, df=12, n.s.). Both were strongly associated with distress, as confirmed by strong Cramer's V correlations (0.76).



Figure 4. Acquisition of reading skills in Japanese and English (n = 10)

The vertical axis is Languages; Hiragana, Katakana, and Kanji are Japanese. The horizontal axis is Answers. □unable to read, ■able to read.



Figure 5. Acquisition of writing skills in Japanese and English (n = 10)

The vertical axis is Languages; Hiragana, Katakana, and Kanji are Japanese. The Horizontal axis is Answers. \Box unable to write, \blacksquare able to write.

Table 3. Number of participants who experienced distress while learning languages in school (n = 10)

Responses	Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Mean	SD
Japanese (n)	1	4	2	1	2	3.10	1.37
English (n)	2	3	3	2	0	3.50	1.08

Discussion

The present study examined 3D letter expressions for children and young adults with dyslexia and their intuitive recognition of letter structure in English learning material on a tablet. The results revealed that stroke order was most easily understood in letter expression type A (black) for both letters "b" and "d." The data obtained suggest that depth and color expressions in 3D letters are key factors in recognizing alphabetical letters efficiently. The depth of information regarding a letter's structure might be important to assist English learning for individuals with dyslexia. Stroke order of alphabetical letters may also be more easily recognized if letter color and background color have a high contrast ratio; letter expression type E (Gradation 1) was easily understood for "d" than type C (Gradation 1). Further, easy understanding of gradation of 3D letters.

Children with dyslexia present difficulty learning English; our results are consistent with those of Wydell and Butterworth's (1999) study, which examined the difficulty of learning English in school. We found participants struggled to learn English if they also struggled to learn Kanji. Kanji has different letter structures than Hiragana and Katakana, as well

as more strokes; participants were able to read Kanji but could not recall the order when writing it (Uno et al., 2009). Further, Kanji has ideographic characters, making their meaning easy to understand; hence, reading is not difficult. In the case of English, meaning-making is not possible based on the shape of its letters; hence, learning to read and write English is likely to be more difficult for Japanese learners.

Our survey results on the relationship between Japanese and English acquisition suggest that if students find it difficult to read and write Kanji in elementary and junior high school, it might also be difficult for them to learn English in junior high school. In Japanese education, English classes begin in the upper grades of elementary school, but the number of hours of English increases in junior high school, making the subject more difficult. Our results suggest that children who have difficulty acquiring Kanji, especially writing Kanji, also struggle to learn English. The early detection of children presenting difficulty in Kanji might lead to earlier support for English learning.

We found that half the participants experienced distress while learning languages in school. The most-used method for learning Japanese requires learners to write Kanji characters physically while looking at a model. Learners need to write the characters repeatedly until they master writing them correctly by themselves. In Japanese education, stroke order is important. This method has been adopted in the learning of English, as well. According to participants' responses, writing repetition and maintaining stroke order is very spiritual. Distress while learning languages in school may not only affect learning native and foreign languages but may also cause a loss of motivation for other learning. Educators must determine whether a child requires repeated learning for acquiring native or foreign languages.

Our study has not fully clarified our teaching materials' effect on learning English. In the future, once their effect has been clarified, our method may provide an easy learning strategy as compared to conventional iterative learning and may lead to long-term language retention.

Japanese children feel resistance toward learning English, regardless of dyslexia (Ministry of Education, Culture, Sports, Science and Technology, 2004). Therefore, more research is needed that examines all children in general. Since these aspects could not be clarified in this study, they should be focused on in future research. Learning a language requires a method that provides children enjoyment rather than struggle.

Conclusion

Recently, electronic devices have become widespread and have been applied in English learning environments. In this study, we examined 3D letters as a novel learning approach. The results point to the possibility that differences in letter expression may affect the ease of understanding letters' stroke order. A limitation of this study is its small sample size; future studies should examine the influence of 3D letters on English learning capabilities with a larger sample of children with dyslexia. The guestionnaire-survey was conducted among children with dyslexia regarding difficulties in learning their native language and foreign language in school. The results indicate that difficulties in acquiring Japanese are also reflected in learning English in children with dyslexia. In addition, it might be necessary for teachers to consider the distress caused by English learning in children with dyslexia. If a child has difficulties in writing Kanji, it is quite possible they will also have trouble learning English; hence, learning support from elementary school to junior high school is important. The opportunity to learn should not be lost due to the distress of learning languages in school. Relieving this distress may be important for successful language learning. A method incorporating 3D technology may help alleviate this distress in children with dyslexia. In the future, we plan to investigate the effects of 2D and 3D learning techniques and potential distress during 3D learning. We also plan to investigate the detailed effect of 3D learning techniques on English reading and writing.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP 17K18195.

References

- Ferrer, E., Shaywitz, B. A., Holahan, J. M., Marchione, K., & Shaywitz, S. E. (2010). Uncoupling of reading and IQ over time: Empirical evidence for a definition of dyslexia. *Psychological Science*, 21(1), 93–101. https://doi.org/10.1177/0956797609354084
- Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1992). A twin study of the etiology of comorbidity: Attention-deficit hyperactivity disorder and dyslexia. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*(2), 343–348. https://doi.org/10.1097/00004583-199203000-00024
- Knoop-van Campen, C. A. N., Segers, E., & Verhoeven, L. (2020). Effects of audio support on multimedia learning processes and outcomes in students with dyslexia. *Computers and Education*, 150, 103858. https://doi.org/10.1016/j.compedu.2020.103858
- Kooi, F. L., & Toet, A. (2004). Visual comfort of binocular and 3D displays. *Displays*, 25(2–3), 99–108. https://doi.org/10.1016/j.displa.2004.07.004
- Kronenberger, W. G., & Dunn, D. W. (2003). Learning disorders. *Neurologic Clinics*, 21(4), 941–952. https://doi.org/10.1016/S0733-8619(03)00010-0
- Lambooij, M., Fortuin, M., Heynderickx, I., & Ijsselsteijn, W. (2009). Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science* and Technology, 53(3), 30201-1.
 - https://doi.org/10.2352/J.ImagingSci.Technol.2009.53.3.030201
- Lyon, G. R., Shaywitz, S. E., & Shaywitz, B. A. (2003). A definition of dyslexia. *Annals of Dyslexia*, 53(1), 1–14. https://doi.org/10.1007/s11881-003-0001-9
- Ministry of Education, Culture, Sports, Science and Technology. (2004). Shogakkou no Eigo Kyoiku ni Kansuru Ishiki Chousa Chousa Houkoku-sho [Awareness survey report on English education in elementary schools]. https://dl.ndl.go.jp/info:ndljp/pid/1206332
- Pastor, P. N., & Reuben, C. A. (2008). Diagnosed attention deficit hyperactivity disorder and learning disability: United States, 2004–2006. National Center for Health Statistics. Vital and Health Statistics. 10(237). https://www.cdc.gov/nchs/data/series/sr_10/sr10_237.pdf

Read, J. C. A., & Bohr, I. (2014). User experience while viewing stereoscopic 3D television. *Ergonomics*, *57*(8),1140–1153. https://doi.org/10.1080/00140139.2014.914581

- Thurstone, L. L. (1927). Psychophysical analysis. American Journal of Psychology, 38(3), 368–389. https://doi.org/10.2307/1415006
- Uno, A., Wydell, T. N., Haruhara, N., Kaneko, M., & Shinya, N. (2009). Relationship between reading/writing skills and cognitive abilities among Japanese primary-school children: Normal readers versus poor readers (dyslexics). *Reading and Writing*, 22(7), 755–789. https://doi.org/10.1007/s11145-008-9128-8
- von Károlyi, C., Winner, E., Gray, W., & Sherman, G. F. (2003). Dyslexia linked to talent: Global visual-spatial ability. *Brain and Language*, *85*(3), 427–431. https://doi.org/10.1016/s0093-934x(03)00052-x
- Wydell, T. N., & Butterworth, B. (1999). A case study of an English-Japanese bilingual with monolingual dyslexia. *Cognition*, *70*(3), 273–305. https://doi.org/10.1016/s0010-0277(99)00016-5
- Yamazoe. (Ikeshita), H., Mitsuya, R. Kawai, T., Sato, T., Yamagata, H., Yamazaki, T., & Miyao, M. (2009). The effect of using stereoscopic in a literacy learning system for children with developmental dyslexia. *Educational Technology Research*, 32(1–2), 53–60.
Hsin-chou Huang, Barbara Loranc

National Taiwan Ocean University, Keelung, Taiwan University of Bielsko-Biała, Bielsko-Biała, Poland

joehuang@ntou.edu.tw, bloranc@gmail.com

EFL learners' oral communication strategies: Insights from a Taiwan and Poland exchange project

Bio data



Hsin-chou Huang is Professor of English as a Foreign Language in the Institute of Applied English at National Taiwan Ocean University. Her research interests include CALL, second language reading and writing, and telecollaboration for intercultural learning. She has published on these topics in scholarly journals such as Computer-Assisted Language Learning, ReCALL, Computers & Education, and Language Teaching Research.

Barbara Loranc is Assistant Professor at University of Bielsko-Biała, Poland. She holds a PhD in applied linguistics from the University of Silesia. Her research interests focus on various aspects of telecollaboration for pre-service teacher training and innovative uses of digital resources in foreign language teaching.

She has published in international journals and edited volumes on second language acquisition.

Abstract

Recent studies have confirmed the benefits of virtual exchanges for enhancing EFL learners' language skills, digital literacy and intercultural competence (O'Dowd & O'Rourke, 2019). This study examined oral communication strategies to determine which proved most effective for EFL learners attempting to communicate with global peers. Participants were 32 non-English majors from a national university in northern Taiwan and 38 education majors from a public university in Poland. During the 12-week exchange project, students engaged in three types of interactions: information exchange tasks to share their understanding of given cultural topics, comparison and analysis tasks to foster critical thinking about the topics, and product creation tasks, making use of technological tools to create videos showing students' interpretations of cultural issues. Data from an oral communication survey revealed that students relied on various strategies to cope with speaking and listening problems. To deal with speaking problems, students gave examples if a listener did not understand. They also used gestures and facial expressions to facilitate communication. With regard to listening problems, students often guessed a speaker's intention based on identification of familiar words and asking for repetition. Students appreciated the chance to have authentic encounters with global peers with whom they could use language for communication purposes. They learned about foreign cultures and gained greater respect for cultural differences. This study's findings on students' use of oral communication strategies and degrees of enhancement in intercultural competence will help future teachers determine how to best engage students in telecollaborative projects.

Conference paper

Introduction

Telecollaboration, also known as virtual exchange or online international exchange, involves "internet-based intercultural exchange between people of different cultural/national backgrounds, set up in an institutional context with the aim of developing both language skills and intercultural communicative competence through structured tasks" (Guth & Helm, 2010, p. 14). As a result of globalization and the rise of digital technologies, telecollaboration has gained popularity with teachers and researchers as an effective and powerful means to enhance second language acquisition and foster intercultural communication competence (Cunningham, 2019; Dooly & O'Dowd, 2018; Goodwin-Jones, 2019; O'Dowd, 2018; O'Dowd & O'Rourke, 2019).

The use of videoconferencing to foster authentic interactions among global peers is helpful in improving learners' command of target languages (Bruun, 2018; Kato, Spring, & Mori, 2016), enhancing their use of semiotic modes in meaning-making (Dooly & Hauck, 2012), assisting them negotiate identities (Yang & Yi, 2017), and fostering critical thinking (Semercu & Aydin, 2018). To achieve the benefits of synchronous interactive learning environments, educators nowadays need to combine authentic tasks and digital communication tools in their classrooms. Teachers should motivate their students to use multimodal resources available in digital communication environments as ways to help students develop effective communication strategies (Dooly & Hauck, 2012; Dzekoe, 2017).

Given increasing opportunities for global peers to collaborate, and the easy availability of video conferencing, this study aimed to investigate how global peers work through telecollaborative tasks in a synchronous context. Specifically, the study explored EFL learners' oral communication strategies when working with global peers. Illuminating learners' speaking and listening problems and their strategies to address these problems may help educators design communicative tasks with a clear impact on performance. The research question is set as follows: What are students' oral communication strategies in interacting with global peers, and how do students improve these strategies?

Methodology

This project aimed to uncover oral communication strategies among EFL telecollaborators. The researcher designed a language curriculum with culturally informed activities to immerse students in real communication with global peers. Thirty-two non-English majors from the researcher's freshman English class at a national university in Taiwan were paired with thirty-eight education majors from the University of Bielsko-Biala, Poland. The students from Taiwan had been studying English for at least six years and had an intermediate level of proficiency, while the Polish students were proficient second language users. As the total number of participants from Poland was six more than that from Taiwan, thirty-two pairs of Taiwan-Polish students in a group. The collaboration project between Polish and Taiwanese students provided a chance for both groups to broaden their international experience and enhance their cultural awareness.

The Language Exchange Task

This study utilized O'Dowd and Ware's (2009), and Guth and Helm's (2010) ideas on integrating online literacy skills into the design of CMC-based intercultural tasks, defined as information exchange, comparison and analysis, collaboration, and product creation. During the 12-week exchange project, students engaged in three types of interactions: information exchange tasks to share their understanding of given cultural topics,

comparison and analysis tasks to foster discussion and critical thinking about topics; and product creation tasks to make use of technological production tools to create videos showing students' interpretations of cultural issues. In the first information exchange stage, students got to know one another through cultural exploration. Topics for exchange consisted of (1) general facts about the partner nation, (2) food and culture, (3) impact of COVID-19, and (4) tourist attractions. Exploration of each topic lasted for two weeks. The researcher, in her role as teacher, prepared authentic materials such as online readings or videos for students to review before involving partners in actual communication. Students exchanged cultural topics through synchronous communication in pairs via Zoom.

In the second stage, involving comparison and analysis, students compared similarities and differences about the topics discussed in the previous stage. The researcher prepared prompt questions to facilitate dialogue. In the final stage of product-creation, both Taiwan and Polish students synthesized information to create films or Powerpoints in pairs. Students worked in pairs to prepare one video per pair, with content consisting of perspectives from Taiwan, perspectives from Poland, similarities between Taiwan and Poland, and differences between Taiwan and Poland.

Data Collection and Analysis Procedures

To reveal students' oral communication strategies, a questionnaire adapted from Nakatani's (2006, p.163-164) oral communication inventory was used at the end of the project. This inventory was designed to explore EFL students' speaking and listening strategies, with eight categories dealing with speaking problems and seven categories describing listening problems during communication. The instrument relies on a Likert-type scale with 32 items classified into eight dimensions for speaking: social affective, fluency-oriented, negotiating for meaning while speaking, accuracy-oriented, message reduction and alteration, nonverbal strategies while speaking, message abandonment, and attempts to think in English. Strategies for coping while listening included 26 items classified into seven dimensions: negotiating for meaning while listening, fluency-maintaining, scanning, getting the gist, nonverbal strategies while listening, less active listener, and word-oriented. Students also attended semi-structured interviews at the end of the project to elicit their feedback on oral communication tactics. The oral communication survey was analyzed using descriptive statistics, and gualitative data from semi-structured interviews were coded based on Tesch's (1990) qualitative research.

Results

Overall, students showed competence in using various strategies to cope with speaking and listening problems (M = 3.87, SD = 0.87), with more use of listening strategies (M =3.93, SD = 0.84) than speaking ones (M = 3.83, SD = 0.89) (see Table 1). To deal with speaking problems, students used the strategy of giving examples if a listener did not understand (M = 4.31, SD = 0.77); they also used gestures and facial expressions if they believed they had failed to communicate well (M = 4.19, SD = 0.77). They applied strategies such as paying attention to a listener's reaction (M = 4.19, SD = 0.77) and taking their time to express themselves (M = 4.16, SD = 0.75). When they felt incapable of expressing their original idea, they tried to replace the original message with another one (M = 4.09, SD = 0.72) and used simple expressions (M = 4.03, SD = 0.77). Other strategies included using fillers (M = 3.91, SD = 0.80), making eye contact (M = 3.88, SD = 0.82), and asking for help (M = 3.91, SD = 0.88) when communication broke down.

With regard to listening problems, students often guessed a speaker's intention by picking up familiar words (M = 4.131, SD = 0.68) and asking for repetition when they

could not understand what the speaker said (M = 4.28, SD = 0.76). They also tried to catch the speaker's point and focus on every word that the speaker used (M = 4.09, SD = 0.72). Students paid attention to words that the speaker pronounced slowly or emphasized (M = 4.06, SD = 0.75). They paid attention to the first part of a sentence and guessed the speaker's intention (M = 4.00, SD = 0.75). They asked the speaker the speaker to use easy words when they had difficulties in comprehension (M = 3.91, SD = 0.84) and translated into native language to build a gradual understanding of what the speaker had said(M=3.88, SD=0.99).

Table 1.	Results o	f Oral	Communication	Strategy	Use
----------	-----------	--------	---------------	----------	-----

Rank	Question	Mean	SD
(I) St	rategies for Coping with Speaking Problem	3.83	0.89
1	20. I give examples if the listener doesn't understand what I am	4.31	0.77
	saying.		
2	25. I try to give a good impression to the listener.	4.22	0.74
3	16. I use gestures and facial expressions if I can't express myself	4.19	0.77
4	19 While speaking I hav attention to the listener's reaction to	4 19	0 77
-	my speech.	1.19	0.77
5	3. I use words that are familiar to me.	4.16	0.67
6	10. I take my time to express what I want to say.	4.16	0.75
7	5. I replace the original message with another message when I	4.09	0.72
	believe I have failed to communicate well the first time.		
8	4. I reduce the message and use simple expressions.	4.03	0.77
9	17. I correct myself when I notice that I have made a mistake.	4.00	0.79
10	22. I make comprehension checks to ensure that the listener	4.00	0.83
	understands what I want to say.		
11	29. I actively encourage myself to express what I want to say.	3.97	0.68
12	2. I think first of a sentence I already know in English and then	3.91	0.68
	try to change it to fit the situation.		
13	23. I try to use fillers when I cannot think of what to say.	3.91	0.80
14	26. I don't mind taking risks even though I might make mistakes.	3.91	0.84
15	31. I ask other people to help when I can't communicate well.	3.91	0.88
16	I try to make eye-contact when I am talking.	3.88	0.82
17	1. I think first of what I want to say in my native language and	3.78	0.89
10	then construct the English sentence.	2 70	0.00
18	11. I pay attention to my pronunciation.	3.78	0.82
19	14. I pay attention to the conversation flow.	3.78	0.82
20	27. I try to enjoy the conversation.	3.78	0.93
21	12. I try to speak clearly and loudly to make myself heard.	3.75	0.75
22	I pay attention to my rhythm and intonation.	3.75	0.79
23	30. I try to talk like a native speaker.	3.75	0.97
24	18. I notice myself using an expression that fits a rule I have learned.	3.66	0.73
25	9. I change my way of saving things according to the context.	3.59	0.82
26	21 I repeat what I want to say until the listener understands	3 59	0.86
20	28. I try to relay when I feel anyious	3 50	1 00
27	8. I try to emphasize the subject and yerb of the centence	3 56	0.07
20	6. I abandon the execution of a verbal plan and just say some	2.20	0.97
29	words when I don't know what to say.	5.55	0.97
30	24. I leave a message unfinished because of some language	3.50	1.03
	difficulty.		
31	7. I pay attention to grammar and word order during conversation.	3.41	0.86
32	32. I give up when I can't make myself understood.	2.78	1.05

(II) S	trategies for Coping with Listening Problems	3.93	0.84
1	3. I guess the speaker's intention by picking up familiar words.	4.31	0.68
2	22. I ask for repetition when I can't understand what the speaker	4.28	0.76
	has said.		
3	I try to catch the speaker's main point.	4.19	0.77
4	I try to catch every word that the speaker uses.	4.09	0.72
5	6. I try to respond to the speaker even when I don't understand	4.09	0.84
	him/her perfectly.		
6	7. I guess the speaker's intention based on what he/she has said	4.09	0.76
	so far.		
7	4. I pay attention to the words that the speaker pronounces	4.06	0.75
_	slowly or emphasizes.		
8	18. I pay attention to the speaker's eye contact, facial expression	4.06	0.75
•	and gestures.		
9	1. I pay attention to the first word to judge whether it is an	4.03	0.//
10	Interrogative sentence or not.	4 00	0 75
10	5. I pay attention to the first part of the sentence and guess the	4.00	0.75
11	Speaker's intention.	4 00	0.70
11	I don't understand his/her intention well	4.00	0.79
12	Q I anticipate what the speaker is going to say based on the	3 07	0 73
12	ontext	5.97	0.75
13	14 I send continuation signals to show my understanding in	3 94	0 75
15	order to avoid communication gaps	5.54	0.75
14	21. I make a clarification request when I am not sure what the	3.94	0.86
	speaker has said.	0.5.	0.00
15	25. I especially pay attention to the interrogative when I listen to	3.94	0.90
	WH-questions.		
16	20. I ask the speaker to use easy words when I have difficulties	3.91	0.84
	in comprehension.		
17	11. I try to translate into native language little by little to	3.88	0.99
	understand what the speaker has said.		
18	I pay attention to the speaker's pronunciation.	3.88	0.74
19	I pay attention to the speaker's rhythm and intonation.	3.84	0.75
20	17. I use gestures when I have difficulties in understanding.	3.84	0.87
21	10. I ask the speaker to give an example when I am not sure	3.81	0.88
	what he/she said.		
22	26. I pay attention to the subject and verb of the sentence when	3.72	0.84
~~	I listen.		
23	19. I ask the speaker to slow down when I can't understand what	3.66	0.96
24	the speaker has said.	2.66	0.01
24	23. I make clear to the speaker what I haven't been able to	3.66	0.81
25	understand.	2 50	1.00
20 26	o. I don't minu il I can't understand every single detall.	3.39	1.03
Z0	24. I Only TOCUS ON TAINING EXPLESSIONS.	2.44	0.79
Total		3.07	0.87

Note: This questionnaire was adapted from Nakatani's (2006 pp. 163-164) oral communication inventory.

Semi-structured interview results supplemented findings from the oral communication survey. When students were asked about their listening problems in Zoom sessions, they indicated that they were overwhelmed by the fast speed of their Polish peers' speaking and their accents. Students commented:

It was sometimes difficult when she was speaking too fast or using too complicated words, which prevented me from understanding my partner. (Student #10)

68

This was my first time to communicate with a European, so I had a hard time understanding what he said because of the accent. (Student #20)

When asked about the strategies they used to cope with challenges, students reported that they asked their peers to repeat, or asked them to send text messages when needed. Other strategies included asking for help from a friend, or using Internet resources to illustrate what they wanted to say. As students worked with the same partner, some chose to listen to the recorded meeting again, trying to guess meanings they had initially missed based on the context and noting down the words and phrases that the partner used so that they could better understand in the next meeting. Students also used strategies such as finding synonyms for difficult words, changing the topic, and using body language to keep the conversation going. Students commented:

When needed, both sides could ask the other to repeat what was said, without any awkwardness involved. (Student #3)

So, I listened to the recording to let myself become familiar with the accent and speaking speed of my peer. Now I can understand more easily what she says. (Student #8)

We looked for synonyms of the words we meant or tried to break sentences down to much simpler English. (Student #11)

With regards to speaking problems, students made clear that their lack of vocabulary and grammar knowledge lowered their confidence in speaking. Students commented:

Sometimes we had speaking problems. I couldn't express my ideas with the right words, so I tried to get meaning across by using body language or sharing a photo with my partner so she could more easily understand a word or sentence. (Student #15)

It was hard for me to organize my sentences in a short period of time. Also, I was not always sure whether the word I used was correct or not. Fortunately, my partner was good at guessing what I said even though I didn't always know what I was talking about. (Student #22)

When asked about strategies for coping with speaking problems, students said that they used realia at hand when doing video conferencing, such as showing pictures of a word they wanted to say. Alternatively, students used body language or texted their peers in the chatroom if they could not pronounce a word understandably.

When difficulties arose, I would try to use another method to replace "talking". For example, I would show pictures or videos online that could put meaning across successfully. (Student #2)

Besides communicating by talking to each other, sometimes when I didn't know how to explain an idea, I would use body language to let my partner know what I meant. (Student #21)

When asked what they had learned from communicating with global peers, they indicated that they had learned strategies to boost their confidence when speaking to foreigners. They also learned to respect cultural differences. From the actual final collaborative project, they learned teamwork and how to work with foreign peers in a more efficient way.

I think *I* learned many things after interacting with Polish students. *I* was very nervous at first. Now we can chat easily and finish the final project together. *I*

Proceedings of the XXIst International CALL Research Conference

69

learned many things about Poland and its food, culture, architecture, transportation and so on. (Student #25)

Discussion and Conclusion

This project aimed to engage EFL students in authentic language learning experiences through telecollaborating with global peers to complete tasks. During synchronous Zoom interactions, students succeeded in enhancing oral communication strategies (Dezkoe, 2017; Dooly & Hauck, 2012). With the interactive features of video conferencing, students received instant verbal and non-verbal feedback from interlocutors (Kato et al., 2016). When encountering difficulties in expressing themselves, they used a wide range of communicative resources, such as simplifying expressions and attempting to think in English (Nakatani, 2006). Students also applied non-verbal strategies (gestures and facial expressions) to convey meaning (Austin et al., 2017). The use of multimodal aids like videos and pictures supplemented their efforts to deliver key concepts to global peers. The use of realia lowered students' emotional barriers while they were speaking and listening to foreign counterparts (Sanchez & Manrique, 2018). This telecollaborative project enabled students to become aware of their own use of oral communication strategies and critically reflect on their strengths and weaknesses in terms of linguistic resources during synchronous chats. The end result was a more sophisticated meaning exchange.

In an increasingly interconnected world, this study shed light on how EFL students can benefit from telecollaborative projects via structured virtual exchange tasks. Findings on students' dyadic interactions in video conferencing and on effective oral strategies may improve future telecollaborative projects and allow teachers and researchers to develop more efficient methods of imparting competence. Communication capabilities are essential for those who must collaborate and negotiate meaning during online interactions with global peers in the digital age.

References

- Austin, N., Hampel, R., & Kukulska-Hulme, A. (2017). Video conferencing and multimodal expression of voice: Children's conversations using Skype for second language development in a telecollaborative setting. *System*, 64, 87-103.
- Bruun, S. (2018). Global goals: A virtual project with students from Sweden and Tanzania. In M. Dooly & R. O'Dowd (Eds.), In this together: Teachers' experiences with transnational, telecollaborative language learning projects (pp. 199-215.) Peter Lang.
- Cunningham, D. J. (2019). Telecollaboration for content and language learning: A genre-based approach. *Language Learning & Technology*, 23(3), 161–177.
- Dooly, M., & Hauck, M. (2012). Researching multimodal communicative competence in video and audio telecollaborative encounters. In M. Dooly & R. O'Dowd (Eds.), *Researching online interaction and exchange in foreign language education* (pp. 135-162.). Peter Lang.
- Dooly, M., & O'Dowd, R. (2018). Telecollaboration in the foreign language classroom: A review of its origins and its application to language teaching practice. In M. Dooly & R. O'Dowd (Eds.), In this together: Teachers' experiences with transnational, telecollaborative language learning projects (pp. 12-34.) Peter Lang.
- Dzekoe, R. (2017). Computer-based multimodal composing activities, self-revision, and L2 acquisition through writing. *Language Learning & Technology*, 21(2), 73–95. http://llt.msu.edu/issues/june2017/dzekoe.pdf
- Godwin-Jones, R. (2019). Telecollaboration as an approach to developing intercultural communication competence. *Language Learning & Technology*, 23(3), 8–28.

- Guth, S., & Helm, F. (2010). *Telecollaboration 2.0: Language, literacies and intercultural learning in the 21st century*. Peter Lang.
- Kato, F., Spring, R., & Mori, C. (2016). Mutually beneficial foreign language learning: Creating meaningful interactions through video-synchronous computer-mediated communication. *Foreign Language Annals*, 49(2), 355-366.
- Nakatani, Y. (2006). Developing an oral communication strategy inventory. *The Modern Language Journal*, *90*, 151-168.
 - https://doi.org/10.1111/j.1540-4781.2006.00390.x
- O'Dowd, R. (2018). From telecollaboration to virtual exchange: State-of-the-art and the role of UNICollaboration in moving forward. *Journal of Virtual Exchange*, 1, 1–23. https://doi.org/10.14705/rpnet.2018.jve.1
- O'Dowd, D., & Ware, P. (2009). Critical issues in telecollaborative task design. *Computer* Assisted Language Learning, 22(2), 173-188.
- O'Dowd, R., & O'Rourke, B. (2019). New developments in virtual exchange in foreign language education. *Language Learning & Technology*, 23(3), 1-7.
- Sanchez, G. B., & Manrique, G. G. (2018). What makes our school unique? A telecollaborative experience from the perspective of two "new-comers". In M. Dooly & R. O'Dowd (Eds.), In this together: Teachers' experiences with transnational, telecollaborative language learning projects (pp. 145-181.) Peter Lang.
- Semerci, A., & Aydin, M. K. (2018). Examining high school teachers' attitudes towards ICT use in education. *International Journal of Progressive Education*, 14(2), 93-105.
- Tesch, R. (1990). Qualitative research: Analysis types and software tools. Faler Press.
- Yang, S. J., & Yi, Y. (2017). Negotiating multiple identities through eTandem learning experiences. *CALICO Journal, 34*(1), 97-114.

Hsin-Yi Cyndi Huang, Chiung-Jung Tseng, Ming-Fen Lo

Center for Bilingual Education, Southern Taiwan University of Science and Technology, Taiwan

cyndihuang@stust.edu.tw, zt7@stust.edu.tw, mingfen@stust.edu.tw

Applying translanguaging pedagogy to scaffold non-English major juniors on writing scripts for English presentations

Bio data



Dr. Hsin-Yi Cyndi Huang is an assistant professor in the Center for Bilingual Education at Southern Taiwan University of Science and Technology (STUST). She received her Ed.D. in TESOL from Alliant International University in 2012. Her research interests include teaching EFL writing, pronunciation teaching, and technology-enhanced language learning and teaching.



Dr. Chiung-Jung Tseng is an assistant professor at the Center for Bilingual Education, Southern Taiwan University of Science and Technology (STUST). She obtained her PhD in Literature from National Cheng Kung University in 2011. Her research interests include nineteenth-century literature, Gothic fiction and English language learning and teaching.



Dr. Ming-Fen Lo is a Lecturer in the Center for Bilingual Education at Southern Taiwan University of Science & Technology (STUST), Tainan, Taiwan. She received her PhD in Educational and Applied Linguistics from University of Newcastle upon Tyne in 2017. Dr. Lo's main research interests include Conversation Analysis, L2 classroom interaction, and technology-assisted language learning and teaching.

Abstract

To tackle students' English writing problems while they are drafting scripts for presentations, in this project, we propose applying translanguaging pedagogy as well as smartly utilizing Google Translate to scaffold college juniors on writing scripts for their presentations. The participants are from three intact classes of non-English major juniors taking the required "Advanced English Expression" course, where each participant will draft their scripts in Chinese first, then translate them both on their own and through Google Translate. During the revision period, teacher's consultation will also be offered. By referring to their Google-translated version text, the students revise their self-written English scripts to achieve what they want to express. We believe that through such curriculum planning and design, the students can learn to make good use of their mother tongue and technology, cultivate critical thinking skills, and improve their knowledge and skills in English writing. This study is based on an action research method and goes to great lengths to increase teaching quality in the long run. Three versions (self-written version, Google translated version and post-edited version) of scripts from each participant in the three intact classes will be collected and further divided into two groups according to the students' English proficiency levels for analysis. Two online writing assessment software (VocabProfiler and Scribens) are used for quantitative analysis to compare the differences in grammar and vocabulary of the three texts. Also, three trained raters employ qualitative text analysis by closely examining any improved changes in grammar and word use from each participant's self-written to post-edited versions of scripts. Both the quantitative and qualitative results will be triangulated with the ones from the questionnaires and interviews to gain further insights. It is hoped that this research will shed light on how translanguaging pedagogy can scaffold EFL students on their writing skills and how the students of two different proficiency levels perceive toward this practice. The results of the study can be a great reference to teachers who teach ESL/ EFL writing.

Conference paper

Introduction

Among the four skills of English, writing is often considered the most difficult skill to acquire, especially for low-proficiency EFL learners at universities of technology in Taiwan, Since vocational students, if not English majors, didn't need to prepare for writing tests in the general subject of English for their college entrance exam, they got very limited practice and experience of writing in English during their vocational high school studies. With scant knowledge of grammar and vocabulary, when they are asked to write in English in college, they often directly insert inadequate words as well as grammarly incorrect sentences from online translation without checking the usage, which leads to ineffective communication. It is also worth noting that in a survey conducted in Taiwan, among the 17,000 university participants from universities of technology, only 20 percent of them passed the elementary level of the General English Proficiency Test (GEPT), which is equivalent to A2 level in the Common European Framework of Reference for Language (CEFR) (Wu & Liao, 2003). Nonetheless, in today's era of information technology, writing has become one of the most used forms in communication. Therefore, how to assist the EFL students, especially the low-proficiency students to overcome the English writing barrier has become an urgent issue to tackle.

Notwithstanding the ongoing debate on using students' first language in the English classroom, the smart way of using it might be a light at the end of the tunnel. Even though many still believe that maximizing the input of the target language provides the optimal environment for the English learners during the instruction, advocates and scholars of translanguaging remind us of the goal when teaching and learning English as a foreign language (EFL). Li (2017) pointed out that it appears the goal of teaching and learning EFL is not to get the learners to become another monolingual speaker in English only or to replace their first language, but to make them become bilinguals or even multilinguals. As early as the 1980s, psycholinguist Francois Grosjean (1989) stated that bilinguals are not two monolinguals in one person. Recent brain science studies have revealed that the bilingual or multilingual brain integrates elements of different languages together, and to reach communicative purposes it coordinates, activates, and selects certain parts of the bilingual or multilingual repertoire (Kroll & De Groot, 2009 as cited in Li, 2017). As Canagarajah (2011) noted, for multilinguals, languages are not discrete and separated parts in their repertoires but are an integrated system. For their communicative purposes, multilingual competence, which does not consist of separate competencies for each language, emerges, and this multicompetence functions symbiotically for the different languages in a multilingual's repertoire. For instance, a bilingual is able to use the appropriate language when and with whom. Also, multilinguals are capable of using other resources like paralanguage such as gesture, emotion, etc. for communication (Li, 2017), and this practice is called translanguaging.

Researchers in second language acquisition have found the value in translanguaging to scaffold second language learning for decades (Daniel et al., 2017). Nonetheless, most of the studies were conducted with bilingual learners in an ESL context. Expanding

translanguaging pedagogy to a wider population, namely EFL learners, and investigating its effect in improving EFL writing are worth researching. Thus, the current context of this study can be summarized as follows: the demand for applying translanguaging pedagogy in promoting EFL students' writing skills is increasing, but the reliability of how to smartly incorporate machine translation (MT) like Google Translate (GT) as a tool in the writing process has not yet been fully investigated. We need to determine, for example, how best to use GT, a handy and useful tool with imperfections, to produce the best learning outcomes in students' English writing. Up to now, few studies regarding this issue have been conducted. Hence, it is pivotal to investigate how GT can be incorporated in translanguaging pedagogy to facilitate students' writing process based on empirical evidence and acknowledge its potential benefits, as well as its limitations, in practice. Since most studies regarding L2 writing have been conducted before Google launched its new version of AI-based Google Translate in November 2016, which offers better, more natural translation in whole-sentence translation (Tsai, 2020), consequently, these studies (e.g., Niño, 2004, 2009) reported MT as a bad reference with many lexico-grammatical errors which require much post-editing effort. Two recent significant studies conducted by Lee (2020) and Tsai (2020), respectively, examined the impact of using GT in EFL students' writing. Both studies had the students first write their draft in their L1 and later translate it without the help of GT. Then, the students revised the draft by referring to the GT translated counterparts. In Lee's study, she employed text analysis to examine and compare the differences between the students' first English version and their final version, whereas in Tsai's study he utilized the automated writing evaluation software, VocabProfrofiler (VP) and 1 Checker, to evaluate students' writing performance for the three versions of texts (self-written, Google-translated, and revised self-written). They both found revised texts showed a significant improvement over the self-written texts and concluded MT such as GT could be a useful aid to English writing provided that the instructors can quide the students to be aware of its limitations and offer adequate guidance. This research, on the other hand, does not merely explore the three versions of texts written by two groups (high-and low-proficiency) of EFL non-English majors in university but also investigates whether the inclusion of the instructors or teaching assistants' (TA) consultation in the students' revising process is necessary. In addition, this study employed both the automated writing evaluation software, VP and Scribens, and the text analysis to evaluate the participants' writing performance.

Literature Review

Lee (2020) employed the design where students translated their L1 writing into L2 both with and without the help of MT and then they corrected their L2 writing by referring to the MT translation. She adopted text analysis to analyze both versions of students' writing. The results indicated that using MT helped students correct lexico-grammatical errors in their writing so students' final versions improved significantly in vocabulary, grammar and expressions. In addition, the study also found MT had a positive impact on student writing strategies while the students were editing with MT. She concluded that MT could be a useful aid to language learning; nonetheless, the instructors should make the students aware of its drawbacks and limitations and offer proper guidance while letting students use it. Tsai (2020) investigates the effectiveness of using GT in EFL writing on the university students of both English majors and non-English majors. The participants were asked to write a reflection about a movie in Chinese first, and then their reflection texts were translated both by themselves and by using GT. Finally, they revised their self-translated texts by referring to their GT texts. The analysis of two online computational assessments revealed that, compared with self-translated texts, GT texts showed significantly better writing performance in content, vocabulary and grammar. The revised texts also indicated a significant improvement over self-translated versions, which was especially evident for the non-English major students, who showed significantly positive attitudes toward the use of GT. Moreover, the study suggested a teaching implication of using GT as a revision tool in EFL writing. Although previous studies have found educational benefits of using GT, more research needs to be done on several related issues, like the bigger sample size, different proficiency levels, and so on (Lee, 2020). Therefore, this study investigates GT as a translingual tool in facilitating EFL learners of non-English majors to write their scripts for English presentations. The current study is unique in that it will adopt both text analysis used in Lee (2020) and the automated online computational assessment tools, VocabProfiler and Scribens, to measure how much the students' L2 writing improve across three versions of the scripts with the help of GT and teaching assistants and teachers' consultation. The reason why we incorporate teaching assistants and teachers' consultation into the students' script-rewriting stage is because half of the participants' English proficiency is around A2 in CEFR. Therefore, it could be difficult for them to rewrite and revise the self-written scripts by referring to GT translation scripts on their own. In addition, the participants differences between the two groups on their perceptions toward utilizing GT in English writing. Hence, the following research questions were raised:

- 1. Is there any significant difference among the three versions of scripts (self-translated version, Google Translate version, and post-edited version), written by the participants in terms of writing quality and lexical features?
- 2. Is there any significant difference between the scripts written by the higher proficiency group and the lower proficiency group in terms of writing quality and lexical features?
- 3. What are students' overall perceptions toward using GT as a translingual tool as well as their consultation with the instructor or TA during the revising stage?

Method

Participants and task description

Data for the present study were collected over ten weeks from three intact classes of the same course, English for Professional Communication and Presentation, at a university of technology in southern Taiwan. Each class met for two hours per week over one semester, with the objective of preparing the students with English presentation skills for future careers. The total 146 participants with the same L1, Mandarin Chinese, from three classes were all non-English majors. Their English proficiency levels range between A2 to B1 in CEFR (between high-beginning to intermediate). In order to investigate the effectiveness of using GT to facilitate L2 writing, which may differ based on students' proficiency levels, we gave students a pre-test at the beginning of the semester to divide them into a higher-proficiency (HP) group and a lower-proficiency (LP) group. The HP group comprises 74 students while the LP group consists of 72 students. Since in this course, all the students had to do a 3-minute oral presentation on their chosen topics as their final evaluation, they had to write out the scripts for their presentation first. Therefore, the script-writing activity will involve GT as the translingual tool to facilitate the students' writing and revising process in the current study, which comprised three steps and lasted for nine weeks: Step 1: After the students were instructed with the basic structure and organization of the introductory paragraph for their presentation, in week 3, the students first wrote their introduction in Chinese in class and uploaded their Chinese draft to the school's learning platform, the FlipClass, as their homework. Then, the TA of each class would first check the overall quality of each student's Chinese introduction based on the content and meaning of it. Step 2: In week 4, the students translated in class their Chinese introduction into English without the help of GT or other online dictionaries, and this version is referred to as the self-written (SW) version. Also, the students were asked to translate their Chinese introduction into English solely using GT as the GT version as their homework. Step 3: In week 5, the students edited and revised in class their SW version by comparing it with the GT version. During this week in class and after class, the students can consult with the instructor or TA through LINE to ask questions which they encountered in the editing and revising process. Students were also allowed to use other online resources to help them revise their SW versions. For homework, the students had to upload the revised draft as the post-edited (PE) version to the FlipClass as well. In the following six weeks, the three steps of script-writing activity were repeated for two rounds for writing up the body and conclusion parts of the scripts. Then, the completion of the students' presentation scripts of the three versions (SW, GT, and PE) were collected for further text analysis.

Data collection and analysis

This study employed both the automatic writing evaluation software and manual qualitative text analysis to compare the differences among the three versions of the students' scripts. In addition, data collected from interviews and a survey questionnaire were analyzed both qualitatively and quantitatively as well. All versions of the students' scripts were first analyzed using two types of online assessment freeware, http://www.lextutor.ca/vp/eng/ VocabProfiler(VP) and Scirbens (https://www.scribens.com/). The writing parameters measured by VP include the number of words in the text and the percentage of word appearances in the text of four different categories (K1=the most frequent 1000 words, K2=the second most frequent 1000 words, AWL=academic word list, and off-list words) as well as lexical density. According to Tsai (2020), if the students gain additional information or ideas through the use of their L1, this improvement could be shown by comparing the students' VP writing parameters among different versions of scripts; in addition, these parameters could also reveal vocabulary use. He also noted words in K2 are thought to be more advanced than those in K1. In the current study we adopted the data analysis approach used in Tsai's (2019) study, so the AWL and off-list words will be combined into one category since they are less frequent and involve professional words that students are usually not familiar with. On top of that, lexical density and the numbers and frequencies of different words will also be calculated. The higher the lexical density is, the more information the text is trying to convey. Moreover, grammatical, spelling and punctuation mistakes of the texts could be counted through Scribens, which is a free web-based automated writing assessment tool and detects mistakes in typography & punctuation and grammar immediately. To further compare the differences in writing parameters and errors concerning SW vs. GT, and SW vs. PE scripts, the independent sample t-test and paired sample t-test were employed.

To further investigate the improved changes in grammar and word use from students' SW versions to PE versions, the study also employed qualitative text analysis. Three trained raters examined each participant's SW and PE scripts and identified the improved changes by categorizing them with examples as evidence manually. After the students completed the final scripts of the PE version, a five-point Likert scale questionnaire with 24 questions and six open-ended questions was administered. In all, 18 voluntary participants (nine from the high-proficiency and nine from the low proficiency group, respectively) from each of the three intact classes were interviewed. Each interview took about 15 minutes based on 18 interview questions, which covered the advantages and disadvantages of GT in L2 writing, the help of consultation with the instructor or TA during revision, and the writing strategies they used. Interviews were conducted individually in the instructor's office and recorded. The results of the survey questionnaire and interviews helped to elucidate the students' perceptions toward the use of GT in L2 writing and the grammatical and lexical errors they successfully corrected in their final versions. The quantitative data collected from the questionnaire were analyzed using the EXCEL while the gualitative data gained from open-ended guestions and interviews were coded with multiple steps to identify emerging themes and used to triangulate with the results of script analysis. This study is still ongoing and currently in the phase of data analysis.

References

- Canagarajah, A. S. (2011). Translanguaging in the classroom: Emerging Issues for research and pedagogy. *Applied Linguistics Review*, *2*, 1-28.
- Daniel, S., Jiménez, R., Pray, L., & Pacheco, M. (2017). Scaffolding to make translanguaging a classroom norm. *TESOL Journal*, 10(1), 1-14. https://doi.org/10.1002/tesj.361
- Lee, S.-M. (2020). The impact of using machine translation on EFL students' writing. *Computer Assisted Language Learning*, *33*(3), 157-175.
- Li, W. (2017). *Translanguaging and the goal of TESOL*. Summit on the Future of the TESOL Profession. Retrieved from

- Grosjean, François (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3-15.
- Mehrabiyan, F., & Sharififar, M. (2015). The relationship between translation competence and translator's intelligence. *The Iranian EFL Journal*, *11*(1), 148–163.
- Niño, A. (2004). Recycling MT: A course on foreign language writing via MT post-editing. Paper presented at CLUK (Computational Linguistics United Kingdom 7th Annual Research Colloquium), 6th and 7th January 2004 in the University of Birmingham, United Kingdom. http://www.mt-archive.info/CLUK-2004-Nino.pdf
- Niño, A. (2009). Machine translation in foreign language learning: Language learners' and tutors' perceptions of its advantages and disadvantages. *ReCALL*, 21(2), 241–258. doi:10.1017/S0958344009000172
- Tsai, S. C. (2019). Using Google Translate in EFL drafts: A preliminary investigation. *Computer Assisted Language Learning*, *32*(5), 510-526.
- Tsai, S. C. (2020). Chinese students' perceptions of using Google Translate as a translingual CALL tool in EFL writing. *Computer Assisted Language Learning.* Advance online publication. https://doi:10.1080/09588221.2020.1799412
- Wu, R. H., & Liao, S. Y. (2003). English proficiency test plan for technical college students. Paper presented at the International Applied English Teaching Conference, Ming Chuan University, Taiwan.

https://www.tesol.org/docs/default-source/ppt/li-wei.pdf?sfvrsn=0&sfvrsn=0

Hyangeun Ji, Soyeon Park, Jung Won Jo

Temple University, Philadelphia, United States Hankuk University of Foreign Studies, Seoul, South Korea Hankuk University of Foreign Studies, Seoul, South Korea

hyangeun.ji@temple.edu, cindypark@hufs.ac.kr, jw_jo@hufs.ac.kr

A qualitative study on using Intelligent Personal Assistant in teaching a young Korean learner in an EFL context

Bio data



Hyangeun is a Ph.D. student and research assistant in Science, Math, and Educational Technology at Temple University. Before joining Temple, she worked as a computer programmer and EFL teacher in South Korea. Her research interest includes but is not limited to instructional technology, computer-assisted language learning (CALL), online and virtual learning, and artificial intelligence in education (AIEd).

Soyeon received a BA in English language and literature at Tsinghua University and an MA in TESOL at Hankuk University of Foreign Studies. She also received a Certification of Korean Language Teacher and worked as a Korean language teacher in China. Her research interests include SLA in EFL/KFL contexts, technology-enhanced language learning, online learning, and learners' perception.

Jung Won Jo is an EFL teacher in Korea. She earned her BA in general education at the University of the Cumberlands and is a graduate student at Hankuk University of Foreign Studies. Her research focuses on vocabulary acquisition, reading comprehension, and game-based learning for young language learners.

Abstract

Despite the increasing interest in the potential of Intelligent Personal Assistant (IPA) as a language learning tool, a dearth of empirical studies exists that investigated IPA-based learning for young language learners. To fill this research gap, the current case study explored the use of IPA in teaching a fifth-grade language learner in the Korean EFL context. Specifically, the study aims to explore a language learner's perception of pedagogical benefits of IPA-based learning and attitudes/beliefs about the use of IPA in a language classroom. A semi-structured interview was conducted and analyzed using content analysis. We found that the learner had an overall positive perception about using IPA in class and could identify its pedagogical benefits. The learner's familiarity and interest also increased after the repeated use of IPA. The study provided evidence of integrating IPAs into teaching young language learners and suggests further research on collaboration between IPAs and human teachers.

Conference paper

Innovations in artificial intelligence (AI) technologies brought about changes in the way teachers create learning opportunities and experiences. In particular, advances in natural language processing, machine learning, and automatic speech recognition technologies opened up a new horizon of intelligent personal assistants (IPA)-based learning in and

out of the classroom (Dizon, 2020; Hsu et al., 2021). However, there are still very few empirical studies that examine the use of IPAs in teaching young foreign language learners. Furthermore, the recent emphasis on augmenting teachers' intelligence through collaboration between human teachers and AI (Baker, 2016) has made educational researchers focus on how classroom orchestration could be achieved in an IPA-based learning environment.

Thus, the current case study aimed to fill this research gap by asking two research questions: 1) How do young English as a foreign language (EFL) learners perceive pedagogical benefits of IPA-integrated language learning and 2) What are EFL learners' attitudes and beliefs about using IPAs in language class? A pilot study with six adult EFL learners was conducted, which informed us of the ways to maximize the learning experience in the IPA-based language classroom. Based on the pilot study, we decided to provide the language learner in the current study with guiding questions (e.g., What is the difference between wardrobe and drawer, I would like to know what trouser is) to ask Google Assistant, the IPA used in our study. A fifth-grade female student in Korea from an EFL speaking class participated in the current study. The student received IPA-based language instructions in two 50-minute classes, where learners used Google Assistant to understand the word meaning, check pronunciation, and play an IPA-based word game. After each class, a semi-structured interview was conducted which asked both pedagogical and psychological aspects of using IPAs in class. Using content analysis, the interview data were coded and grouped into pedagogical benefits/challenges and learner's attitudes and beliefs.

The qualitative findings from the interview revealed that the learner's perception of the IPA's pedagogical aspects was mostly positive, while Google Assistant's short recognition time made the learner feel time pressure and anxiety. The learner found the chance to check pronunciation and various resources provided by Google Assistant to understand word meaning particularly valuable. Interestingly, the learner perceived IPA-based language learning as enjoyable and engaging even when faced with a communication breakdown with Google Assistant. The learner added that the teacher's facilitation increased the guality of her interaction with IPA during communication breakdown. Furthermore, the learner was motivated to practice pronunciation to better communicate with Google Assistant. We also found that the learner viewed the IPA-based language learning experience more positively in the second class due to an increased familiarity and interest. As one of the very few attempts that examined the use of IPAs in teaching young EFL learners in a classroom setting, the study proved that IPAs could be efficiently integrated into a language class. Future studies should continue to explore how teachers and AIs could collaboratively create learning spaces for language learners across different age groups and contexts.

References

Baker, R. S. (2016). Stupid tutoring systems, intelligent humans. International Journal of
Artificial Intelligence in Education, 26(2), 600-614.
https://doi.org/10.1007/s40593-016-0105-0
Diron C (2020) Evaluating intelligent percent accistants for L2 listoning and speaking

- Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. *Language Learning & Technology*, 24(1), 16-26. https://doi.org/10125/44705
- Hsu, M. H., Chen, P. S., & Yu, C. S. (2021). Proposing a task-oriented chatbot system for EFL learners speaking practice. *Interactive Learning Environments*, 1-12. https://doi.org/10.1080/10494820.2021.1960864

John Maurice Gayed, May Kristine Jonson Carlon, Jeffrey Scott Cross

Tokyo Institute of Technology, Tokyo, Japan

johnmaurice.gayed@gmail.com, maykristine.jonson@gmail.com, cross.j.aa@m.titech.ac.jp

The Matthew effect in CALL: Examining the equity of a novel intelligent writing assistant as English language support

Bio data



John is a lecturer at the Institute for International Relations teaching students from the School of Engineering at the University of Hyogo. As a Ph.D. student at the Tokyo Institute of Technology (Tokyo Tech), he has been working on AI KAKU, an AI-based writing assistant for English language learners. His other research interests include defining ethical frameworks for the use of AI in education.



May is a postdoctoral researcher at Tokyo Tech, a software engineer at a medium-scale company in Tokyo with a subsidiary in the Philippines, and an Informatics lecturer at Hosei University also in Tokyo, Japan. Her research revolves around improving online learning experiences, may it be through technology-assisted instruction or learning analytics for course quality assurance.



Jeffrey is a professor in Tokyo Tech's School of Environment and Society and the founder of Tokyo Tech's Online Education Development Office (OEDO) – the office develops MOOCs on edX. OEDO also supports Tokyo Tech faculty through the development of online course content, media production, teaching assistant training, and online course learning analytics. In his lab, he supervises students conducting research in educational technology, machine learning, AI in education, and virtual reality.

Abstract

As practitioners introduce new educational technologies into their classrooms, the potential for unintended outcomes from their use might arise. One such potential negative artifact is an increase in the achievement gaps between learners, where high performers tend to benefit more from newly introduced educational technologies than their peers. This phenomenon is commonly referred to as the Matthew effect. In this study, we leverage natural language processing (NLP) based transformers to introduce English language support to English as a Foreign Language (EFL) learners while they are in the writing process. A web-based application was created that uses next-word prediction and automatic reverse translation to help EFL participants in their writing. Adult English language learners from professional development language schools participated in a counterbalanced repeated measures study. To understand the presence of the Matthew effect, learners were grouped based on their self-reported EIKEN scores. Their performance according to two writing factors as well as their perceived cognitive

load while using the tool were measured to establish which groups benefit the most from using the tool. This research sets the stage for understanding how emerging tools can support learning without exacerbating Matthew effects. These effects should be considered in both the development and application of educational technology.

Conference paper

Introduction

The roots of AI in education (AIED) can be traced back to the early 1970s. One of the first initiatives of using AI in the field of education was demonstrated via an intelligent teaching platform called "SCHOLAR CAI" in the United States (Collins & Grignetti, 1975). Since then, the rapid progress of AI technologies has seen many developers and institutions implement these systems with the ultimate goal of making learning and teaching more efficient and effective (Roll & Wylie, 2016).

Recent advancements in natural language processing (NLP) research have brought new opportunities to apply these cutting-edge technologies to computer-assisted language learning (CALL). For instance, grammar and spell check applications have become mainstream tools for English as a Second Language (ESL) / English as a Foreign Language (EFL) educators (Chun et al., 2021; Park, 2019). Thanks to these recent advances in NLP, simple rule-based systems such as grammar checkers have added intelligent context-sensitive features to make the feedback they give users better reflect individual writing styles and intended output. This allows for greater user autonomy and the potential for improved output (Gayed et al., 2022) creating an environment for better learning and learner agency

An issue that CALL practitioners should be aware of is the potential for the Matthew effect to influence the learning outcomes of their students. This effect, for example, can be seen when children fall into different reading levels—stronger readers develop faster and weaker readers fall further behind (Stanovich, 2009). The Matthew effect in language learning can be exacerbated when educational technologies are introduced. The educational technologies due to differences in technology and human support access, making inequalities in education bigger (Reich, 2020). As such, CALL practitioners should be cognizant of which learners are receptive to their interventions, both technology and non-technology related, to prevent disadvantageous positions from being compounded.

This paper focuses on a digital writing assistant and its potential impact on EFL writing. Most current word processing platforms were not built with EFL users in mind and generally give feedback to the user (via grammar and spell-check) only after the user has entered some input into the system. The researchers have developed a digital writing assistant with a basic framework conceptualized around EFLs. Given that this newly developed writing aid has the potential to influence student writing, the researchers explored the equity of using the tool with students with different English skill levels.

Research Questions

This paper examines the intersection of CALL and educational psychology by probing the CALL Matthew effect on the participants of this study. The research questions we are addressing include:

- 1. How much improvement can be detected from different level EFL participants while under experiment and control conditions?
- 2. How prevalent is the CALL Matthew effect among the participants?

Writing proficiency has often been cited as a goal of second language education (Alisaari & Heikkola, 2016) and certainly the goal of language learners themselves. This study introduces a novel digital writing assistant that can potentially aid EFL students in achieving that goal. It is worth noting that even though research has shown smart digital devices have the potential to harm a person's cognitive function (e.g., memory recall) (Tanil & Yong, 2020), we can find little argument for going back to life without smart devices. As such, the removal of smart agents from education is an unpractical approach, yet educators and developers should be more aware of the potential negative impacts smart agents may have on learners.

Related Works

EFL Challenges

There has been much research on the topic of digital tools and their impact on writing. More so, from a CALL perspective, digital mediums have been studied for their possible influence on language learners' ability to write in a second language (L2). Research has shown that writing in a second language is more difficult than writing in one's first language (L1) (Javadi-Safa, 2018; Silva, 1993), and not having strong English writing skills can adversely affect academic performance (Tan, 2011).

A longitudinal study by Laufer (1994) examined the lexical development of advanced second language learners' writing. When the participants' lexical frequency and lexical variation were analyzed, the researcher found only marginal improvements to the former, no improvement in the latter, and no correlation between the two elements was identified. Alfaqiri's (2018) study on Saudi Arabian EFL students investigated the writing difficulties and challenges participants experienced. Data from 114 participants showed that metacognitive strategies were key to improved writing. Additionally, participants' struggle with grammar was identified as a major factor inhibiting higher-level writing production.

Thus, EFL challenges come from at least two fronts: having sufficient lexical and grammatical ability to execute. A common element that restricts L2 writing fluency is the inability to retrieve lexical elements (Schoonen et al., 2009) and having enough cognitive resources to make way for metacognitive strategies that can improve their writing. These two challenges present a feedback loop. For L2 writers, much of the cognitive load comes from translating L1 thoughts to L2 (Nawal, 2018). To be able to think directly in L2 as opposed to translating from L1 and thus optimize cognitive load, writers must have sufficient grammar knowledge and vocabulary to begin with. Retrieving somewhat familiar but not frequently used vocabulary can lead to the tip-of-the-tongue phenomenon which can be frustrating and impede production if not properly resolved (D'Angelo & Humphreys, 2015). To be able to succeed in highly cognitive tasks, one should be able to offload some of the cognitive efforts to the environment whenever practical (Hollan et al., 2000). For L2 writing, being able to produce is arguably more critical than being able to fix grammatical errors, thus these ancillary tasks are good candidates for tool support.

Automated Writing Evaluation

Automated writing evaluation (AWE) systems have gained prominence in digital writing as the sophistication of the feedback available has improved with the integration of NLP technologies. These can be built-in systems (e.g., Microsoft's Editor) or independent software packages (e.g., Grammarly) that can be integrated into existing word processors. AWEs are also slowly becoming popular as language learning support tools. Sevcikova's (2018) study of college-aged participants using AWEs for writing found that the systems can improve language learning. More importantly, students showed greater confidence and motivation while using an AWE. Looking into the accuracy of an AWE and how it compared to human-based assessment, Dodigovic and Tovmasyan (2021) found that the AWE could largely reproduce the quality of human raters when it came to detecting and remediating errors. However, they found certain errors (e.g., coordination, subordination, and relative clauses) were often undetected by AWEs, leading the researchers to the conclusion that AWEs cannot be solely relied upon for evaluation and assessment. Additionally, Zhang's (2020) study on students' use of an AWE showed that engagement with AWEs differed based on the student's English level. Higher-level students were more cognizant of the revision stage of writing and were able to use the feedback they were given more effectively.

CALL Matthew Effect

Confounding factors are commonly exposed and elucidated in second language acquisition research. However, one confounding factor that the researchers found to be less commonly highlighted in CALL literature is the presence and impact of the Matthew effect on learning outcomes (Lamb, 2011). This effect, as seen in Penno et al.'s (2002) study of children's vocabulary acquisition, was seen to be unavoidable across treatment conditions. In the study, treatment interventions were not enough to overcome the effect as higher-level students made greater vocabulary gains than lower-level students. Ngiam and See (2017) examined the link between e-learning CALL applications and music. In their research, the Matthew effect was identified as one negative factor where wealthier students, possessing more cultural capital, were able to perform better than poorer students who did not possess the same level of capital. The poorer students then found themselves in a downward negative spiral, with little awareness of how to improve.

Fortunately, the EFL Matthew effect can be mitigated. For instance, Messer and Nash (2018) were able to minimize the EFL Matthew effect in young English speakers by using visual mnemonics in a CALL study. The researchers found their computer-assisted intervention was effective in improving vocabulary acquisition in the participants. However, as previously mentioned, using the current state-of-the-art AWEs may not be conducive to minimizing the Matthew effect. Even without the usual culprits of the edtech Matthew effect (e.g., technology access and human support), introducing technology can increase the Matthew effect just because the learners do not have the sufficient skill to make sense of the feedback they are given by the technology. We will be referring to the EFL Matthew effect.

Methodology

Treatment Tool - AI KAKU

Advancements in natural language processing and machine learning have led to the development of more sophisticated intelligent writing assistants which offer synchronous feedback to the writer compared to traditional text editors (Frankenberg-Garcia, 2020). In addition, there has been a large volume of research concerning the impact of those digital tools on the writing process (Ashton, 1999; Oh, 2020; O'Regan et al., 2010). AI-assisted writing technology is commonly seen in the form of next-word prediction on smart mobile devices and in some operating systems. Increasingly, next word prediction is becoming a feature available in commonly used word processors such as Google Docs and Microsoft Word. This next-generation type of writing assistance is presented to the user in addition to spelling and grammar correction that users have traditionally experienced. In addition, several applications give further feedback to the user in terms of word suggestions, style feedback, and formative assessment (e.g., Grammarly, Microsoft Editor).

Unfortunately, those tools are primarily aimed at L1 writers and were not intended to assist L2 users with their compositions. Market forces largely dictate software development and there is less demand for digital tools that are intended for the

non-native level English user. This in turn translates to a paucity of literature about the effectiveness of said tools when EFL students are using them. This paper examines a digital writing assistant called "AI KAKU." The name is a take on the Japanese word "書く, kaku," which translates to "to write" in English.

The application was created to assist L2 writers as they are producing written text. The web-accessible artificial intelligence-based writing assistant tool aims to reduce some of the cognitive load that is associated with the second language writing process (Nawal, 2018), allowing users the capability to produce richer, more complex writing than they would without assistance. AI KAKU's interface, as seen in Figure 1, is comprised of five main elements: an input field, a word suggestion engine with confidence scores, a language drop-down menu, a reverse translate output field that translates the users' inputted English into their chosen first language, and a save/export icon for users to be able to download their work.

User input (English only)	AI-KAKU	based next word prediction	
Start writing here こにテキストを入力してください。		Most likely next word is: 最も	可能性の高い
The Edo period in		人 Japan	73 %
		Japanese	5%
		the	5 %
		China	2 %
		modern	1%
Language:言語を選択。 Japanese 期間の期間	Google Translate API		
- 6			

Figure 1. AI KAKU's interface

The framework behind AI KAKU outlined in the previous work of Gayed et al., (2022), will be briefly described here. The next-word prediction is implemented using AllenNLP application programming interface (API) based on Generative Pre-trained Transformer 2 (GPT-2) and the translation is powered by Google Translate API. Only English input is accepted to force thinking in the L2 and default browser grammar and spelling checkers are not blocked. To prevent tool abuse and possible distraction to the writing process, the translation and next-word predictions are only displayed after a 2.5-second delay.

Experimental Design

The researchers utilized a counterbalanced research design with Japanese EFL participants (n = 90) who are studying English at private language schools. The potential effects on student writing while using the AI KAKU application are compared to a control condition without writing assistance. A counterbalanced design minimizes the confounding factors arising from treatment orders and allows all the participants in the study the opportunity to be under the treatment condition. Similar research designs have been employed in L2 research, as seen in Wang's (2019) study of vocabulary recall performance by Chinese students in a university setting or Dizon and Gayed's (2021) study examining Japanese university students using Grammarly as a treatment tool.

The participants were asked to self-report their Test in Practical English Proficiency (EIKEN) scores. The EIKEN test is the most widely used English testing program in

Japan. The exam has a range of seven levels from Grade 5 to Grade 1. Grades 2 and 1 have subgrades (2.5 and 1.5). Grade 1 is the highest-level grade in the exam, being the equivalent of a TOEFL iBT score of 100/120 and Common European Framework of Reference for Languages (CEFR) level C1. Given that our participants are adult learners in optional professional development schools, their economic conditions and adeptness with technology may not be as varied as students in basic education. One way to analyze the equity of educational technology is to compare the performance of low-performing learners with that of high-performing learners (Doroudi & Brunskill, 2019). For this study, the participants were grouped into HIGH (EIKEN 1.5, 2) MIDDLE (EIKEN 2.5), and LOW (EIKEN 3, 4). No participant reported EIKEN level 1 or 5.

After finishing the writing task, the participants were asked to complete a Likert survey that was displayed to the user in both English and Japanese. Perceived usefulness, cognitive load measures, and the number of times word suggestions were used during writing were some of the data points obtained through the survey responses. The participants were randomly assigned to one of four groups as seen in Figure 2.

Group A	Group B	Group C	Group D			
[T] Topic 1	[C] Topic 3	[T] Topic 3	[C] Topic 1			
	Likert	Survey				
[C] Topic 2	[T] Topic 4	[C] Topic 4	[T] Topic 2			
	Likert Survey					
[T] Topic 3	[C] Topic 1	[T] Topic 1	[C] Topic 3			
	Likert Survey					
[C] Topic 4	[T] Topic 2	[C] Topic 2	[T] Topic 4			
Likert Survey						

Figure 2. *Experiment design. T* = *Treatment, C* = *Control*

Lexical Quality Measurements

As for the writing topics the participants were prompted with, four were chosen from a publicly available database of the Test of English as a Foreign Language (TOEFL) administered by Educational Testing Service (ETS). TOEFL is a commonly used English language test administered to foreign students wishing to enter tertiary education in the United States. The researchers chose the Independent Writing Task from the test and all of the questions chosen asked the writer their opinion on commonly discussed social topics. By choosing a standardized test source for our writing prompts, the researchers could avoid weighted difficulty differences between writing prompts. In other words, all of the prompts given to the participants have been validated to be of the same difficulty. The instructions asked participants to write at least three hundred words within the thirty-minute time limit they were given.

To gain objective measurements of writing quality, the researchers used machine assessment to measure three factors. Laufer and Nation's (1995) Lexical Frequency Profile (LFP) examines the word frequencies in a sample text. Less frequent words identified in the British National Corpus (BNC) or the Contemporary American English Corpus (COCA) are considered to be more "advanced" than high-frequency words.

Specifically, the LFP measures the ratio of words written beyond the 2000-word frequency level. Lexical Diversity (LD) is another commonly used measure in second language research. LD identifies the range of different words used in a text. Texts with a lower range tend to use the same words repeatedly, indicating a lack of lexical development and sophistication. LD indices are suggestive of writing quality, vocabulary knowledge, and speaker competence (McCarthy & Jarvis, 2010). Finally, tokens are calculated to measure the rate of production. As an L2 writer progresses in proficiency, their linguistic retrieval speed improves and thus their ability to turn ideas into written text also improves (Palviainen et al., 2012).

Cognitive Load Measurements

Cognitive load, or a person's working memory capacity, is often measured in educational research as a means to gain insight into learning efficiency and efficacy (Clark et al., 2011). This capacity is commonly measured by using offline measurements (e.g., Likert surveys), dual-task measurements (e.g., concurrent load while completing a task), and physiological measurements (e.g., heart rate). Furthermore, cognitive load can be separated into three sub-measurements: intrinsic load, or the relative difficulty of the task at hand; extraneous load, or external load (e.g., noise and distractions) that is caused by elements outside of the problem space; and germane load, or the load associated with the ability to bridge the problem space with existing knowledge.

This study employs offline measurements based on widely used cognitive load rating scales used in educational research. The Paas survey measures overall cognitive load via a nine-point Likert instrument (Paas, 1992). Responses range from 1 [very, very low mental effort] to 9 [very, very high mental effort]. To gain further insight into AI KAKU's potential influence on participants' writing proficiency, the intrinsic load was also measured via a nine-point Likert instrument (Ayres, 2006). Considering one of the researchers' goals while developing AI KAKU was to reduce the problem space for L2 writers, measuring intrinsic load gives the researchers a more granular look into the ability of AI KAKU to address that cognitive burden.

Results and Discussion

Overall Effects

In total, 360 responses were obtained (180 under each writing condition) over the five weeks the study was conducted. After filtering for complete responses, data from 90 respondents were included in this study. Out of the 90 participants, 67 indicated their EIKEN level, data from these participants was used to investigate the CALL Matthew effect. Table 1 shows the breakdown of the respondents according to group assignment, gender, and reported EIKEN levels.

Variables		Values	Dorcontado
variables	Levels	values	reicentage
Group	A	26	28.88%
	В	21	23.33%
	С	20	22.22%
	D	23	25.55%
Gender	Male	34	37.77%
	Female	56	62.22%
EIKEN	1.5	2	2.99%
	2	22	32.84%
	2.5	29	43.28%
	3	13	19.40%
	4	1	1.49%

Table	1.	Demographics	of	participants
-------	----	--------------	----	--------------

Impact of Treatment

Lexical Measures

A paired t-test was used to examine the difference between the control and treatment writing conditions. As seen in Table 2, the measures LFP and LD did not demonstrate statistical significance while the measure of Tokens is significant at p .004, d = 0.2 albeit according to Cohen's d measure, this is conventionally considered a "small" effect size.

To gain more insight into the significant result from the Tokens measure, a scatterplot was plotted, seen in Figure 3, showing the improvement participants demonstrated while under the treatment condition. While under the same writing constraints, the treatment condition allowed participants to produce longer texts, while the lexical diversity and lexical sophistication measures of their writing were largely the same.

Table 2. Lexical differences between writing conditions. Mean and SD values are shown in ()

	Tokens	<i>t</i> -test	LFP	<i>t</i> -test	LD	t-test
Control	156.7 (52.3)	t = -2.8,	0.1 (0.04)	t = -0.19,	61.7 (18)	t = -0.37
Treatment	167.8 (63.2)	df = 179,	0.1 (0.04)	df = 180,	62.2 (18.1	df = 180,
		p = .004		p = .84		p = .7



Means and +/-1 SDs are displayed in red.



Cognitive Load Measures

Since this study takes survey questions out of the Paas (1992) and Ayres (2006) inventory to measure cognitive and intrinsic cognitive load, the researchers needed to confirm the reliability of the questions used in this study. The value for Cronbach Alpha for the survey items was $\alpha = 0.57$, which can be interpreted as "acceptable" according to Taber's (2018) meta-analysis of Alpha reliability measures. Results summarized in Table 3 show that while the difference in overall participant cognitive load did not show statistical significance, the intrinsic load was lower and significant at p .03, d = 0.13; a "small" effect size (Plonsky & Oswald, 2014). A histogram (see Figure 4) of the intrinsic load measure indicates that when participants were writing under the treatment condition, they experienced less perceived difficulty with the writing task at hand.

Proceedings of the XXIst International CALL Research Conference

87

Table 3. Cognitive and intrinsic load differences. Mean and SD values are shown in ().

	Cognitive load	<i>t</i> -test	Intrinsic load	<i>t</i> -test
Control	7.0 (1.4)	t = 0.7, df = 179,	6.3 (1.39)	t = -1.87, df = 179,
Treatment	6.9 (1.3)	p = .4	6.1 (1.48)	p = .03

Note: higher values indicate more load.



Figure 4. Impact of control and treatment on intrinsic load

Two significant outcomes from the experiment show us that participants were able to produce more tokens and felt the inherent difficulty of the writing task was less while they were using the writing assistant (AI KAKU). These results allow the researchers to approach the second research question regarding evidence of the Matthew effect and how the writing assistant impacted participants at different skill levels.

CALL Matthew Effect

As mentioned earlier, participants were grouped into HIGH, MIDDLE, and LOW clusters (n = 67) based on their reported EIKEN levels. To investigate any evidence of the CALL Matthew effect between them, their writing performance and cognitive load measures were examined first across all the EIKEN levels and then across the three levels prescribed by the researchers. The box plots in Figure 5 show the distributions of cognitive load, intrinsic load, lexical frequency, lexical variation, and tokens for each of the assigned EIKEN clusters. The boxplot whiskers extend up to 1.5 * IQR / sqrt(n), where IQR is the interquartile range (the difference between the values at the first quartile and third quartile) and n is the data count. This convention was posited to represent data with a 95% confidence interval when comparing medians for most cases (McGill et al., 1978). Data beyond the whiskers are taken to be the outliers.



Figure 5. Performance per cluster

The figure shows cognitive load decreasing similarly across all three groups; intrinsic load, however, appears to decrease more for the HIGH and MIDDLE clusters, with the LOW cluster experiencing a similar load in both control and treatment conditions. Lexical frequency and lexical variation, interestingly, appear to be negatively influenced by the treatment condition. While the paired *t*-test showed no significance (see Table 2) between control and treatment conditions (EIKEN levels are disregarded here), the researchers feel the results from both lexical frequency and density warrant further investigation. It is possible the AI KAKU writing assistant is introducing additional noise to higher-level participants and somehow hindering or not positively influencing their writing performance. Alternatively, other forms of intervention may be considered to not just improve perceived load but also to affect writing performance more positively.

The researchers decided to split the clusters based on internal discussion and the descriptors of HIGH, MIDDLE and LOW have some flexibility in their definitions (i.e., EIKEN level 2.5 can arguably be considered a "high" level depending on what is being compared). To remove researcher bias in the analysis, a more detailed breakdown of performance per level without clustering can be seen in Figure 6. When broken out of the prescribed clusters, the data suggests higher-level participants are benefitting more from the AI assistant (AI KAKU) than lower-lower participants, suggesting evidence of the Matthew effect. The lexical frequency and diversity for the highest level (EIKEN 1.5) participants clearly show improvement that is not evident at the lower levels.



groups 🛱 control 🛱 treatment

Conclusion and Future Work

90

The data gathered shows evidence that AI KAKU had some positive impact on the L2 writers who participated in this study. The participants produced more words and perceived less mental difficulty when answering the writing prompt with AI KAKU versus without it. While lexical diversity (LD) and lexical sophistication (LFP) did not show any improvement, the researchers believe longer exposure and training with the treatment tool would allow the participants to become more accustomed to the word suggestions and reverse translation provided by AI KAKU. Regardless, the results from this study are promising and further research into AI KAKU is warranted.

Regarding the second research question of evidence of the Matthew effect and how new technology such as AI KAKU impacts users of different skill levels, the researchers could see some effects regarding the cognitive load, lexical frequency, and lexical density. Lower-level users' intrinsic cognitive load remained high despite the assistance AI KAKU gave them during the writing process. On the other hand, higher-level users demonstrate reduced load and improved writing performance while under the treatment condition. Evidence of the CALL Matthew effect in the data supports the argument that higher-level users are benefitting more from the introduced technology than lower-level users. It is to be noted, however, that the distribution of EIKEN levels was heavily skewed to the middle/high levels of 2.5 and 2 and only 3% of the participants reported an EIKEN level is needed to investigate if the effects found in this study can be replicated.

AI KAKU was developed to reduce the cognitive load during the writing process for EFL users. By reducing the problem space and guiding them to think directly in the L2 as opposed to translating their thoughts composed in their L1, learners can hopefully use their cognitive resources on higher-level writing aspects such as organization and revision. An unwanted effect of introducing technology in the learning process, such as in the case of AI KAKU use in English writing, is the widening educational achievement gap or Matthew effect. The researchers recommend instructional designers, CALL developers,

Figure 6. Performance across all EIKEN levels

and in-service educators be more aware of this potentially negative effect of CALL and develop strategies to mitigate the phenomenon.

Further research is needed into these mitigating strategies to reduce the confounding factor of the CALL Matthew effect. The results from this study are in contrast to a similar study by (Chon et al., 2021) that used machine translation (Google Translate) as a mediating agent. The researchers in that study found machine translation assisted the lower-level participants at a greater rate, bringing their performance closer to the higher-level participants. Chon et al's (2021) study does not address the Matthew effect and did not use an explicit mitigating strategy to reduce its effects. A pertinent question is then what are the factors that may exacerbate the Matthew effect among participants.

In addition, further investigation into AI KAKU's impact on the writing process with a wider range of writing quality dimensions, including human assessment of participant writing is warranted. To the same extent that computer-assisted spelling and grammar-check have permeated writing in the modern age, AI-based digital agents will presumably be as commonplace as those older forms of digital assistance. Aspects of their potential should be studied further to ensure equitable access and benefit.

Acknowledgements

This work is supported by the Japan Society for the Promotion of Science (JSPS) via the Grants-in-Aid for Scientific Research (Kakenhi) Grant Number 22K00718 and 20H01719.

References

- Alfaqiri, M. (2018). English second language writing difficulties and challenges among Saudi Arabian language learners. *Journal for the Study of English Linguistics*, 6(1), 24–36.
- Alisaari, J., & Heikkola, L. M. (2016). Increasing fluency in L2 writing with singing. Studies in Second Language Learning and Teaching, 6(2), 271–292.
- Ashton, T. M. (1999). Spell checking: Making writing meaningful in the inclusive classroom. *Teaching Exceptional Children*, *32*(2), 24–27.
- Ayres, P. (2006). Using subjective measures to detect variations of intrinsic cognitive load within problems. *Learning and Instruction*, *16*(5), 389–400.
- Chon, Y. V., Shin, D., & Kim, G. E. (2021). Comparing L2 learners' writing against parallel machine-translated texts: Raters' assessment, linguistic complexity and errors. *System*, 96, 102408.
- Chun, H., Lee, S., & Park, I. (2021). A systematic review of AI technology use in English education. *Multimedia-Assisted Language Learning*, 24(1), 87–103.
- Clark, R. C., Nguyen, F., & Sweller, J. (2011). *Efficiency in learning: Evidence-based guidelines to manage cognitive load*. John Wiley & Sons.
- Collins, A., & Grignetti, M. C. (1975). *Intelligent CAI*. Bolt Beranek and Newman Inc, Cambridge, MA.
- D'Angelo, M. C., & Humphreys, K. R. (2015). Tip-of-the-tongue states reoccur because of implicit learning, but resolving them helps. *Cognition*, 142, 166–190.
- Dizon, G., & Gayed, J. M. (2021). Examining the impact of Grammarly on the quality of mobile L2 writing. *JALT CALL Journal*, *17*(2), 74–92.
- Dodigovic, M., & Tovmasyan, A. (2021). Automated writing evaluation: The accuracy of Grammarly's feedback on form. *International Journal of TESOL Studies*, 3(2), 71–88.
- Doroudi, S., & Brunskill, E. (2019). Fairer but not fair enough on the equitability of knowledge tracing. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 335–339.
- Frankenberg-Garcia, A. (2020). Combining user needs, lexicographic data and digital writing environments. *Language Teaching*, *53*(1), 29–43.

- Gayed, J. M., Carlon, M. K. J., Oriola, A. M., & Cross, J. S. (2022). Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, 100055.
- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. ACM Transactions on Computer-Human Interaction (TOCHI), 7(2), 174–196.
- Javadi-Safa, A. (2018). A brief overview of key issues in second language writing teaching and research. *International Journal of Education and Literacy Studies*, 6(2), 12–25.
- Lamb, M. (2011). A Matthew effect in English language education in a developing country context. In Dreams and realities: Developing countries and the English language (pp. 186–206). The British Council.
- Laufer, B. (1994). The lexical profile of second language writing: Does it change over time?. *RELC Journal*, *25*(2), 21–33.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. Applied Linguistics, 16(3), 307–322.
- McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, 42(2), 381–392.
- McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of box plots. *The American Statistician*, 32(1), 12–16.
- Messer, D., & Nash, G. (2018). An evaluation of the effectiveness of a computer-assisted reading intervention. *Journal of Research in Reading*, 41(1), 140–158.
- Nawal, A. F. (2018). Cognitive load theory in the context of second language academic writing. *Higher Education Pedagogies*, *3*(1), 385–402.
- Ngiam, L. C. W., & See, S. L. (2017). Language e-Learning and music appreciation. In J. I. Kantola, T. Barath, S. Nazir, & T. Andre (Eds.), Advances in human factors, business management, training and education (pp. 865–877). Springer.
- Oh, S. (2020). Second language learners' use of writing resources in writing assessment. *Language Assessment Quarterly*, *17*(1), 60–84.
- O'Regan, B., Mompean, A. R., & Desmet, P. (2010). From spell, grammar and style checkers to writing aids for English and French as a foreign language: Challenges and opportunities. *Revue Francaise de Linguistique Appliquee*, *15*(2), 67–84.
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429.
- Palviainen, Å., Kalaja, P., & Mäntylä, K. (2012). Development of L2 writing: Fluency and proficiency. *AFinLA-e: Soveltavan Kielitieteen Tutkimuksia*, *4*, 47–59.
- Park, J. (2019). An AI-based English grammar checker vs. Human raters in evaluating EFL learners' writing. *Multimedia-Assisted Language Learning*, 22(1), 112–131.
- Penno, J. F., Wilkinson, I. A., & Moore, D. W. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: Do they overcome the Matthew effect? *Journal of Educational Psychology*, 94(1), 23.
- Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, *64*(4), 878–912.
- Reich, J. (2020). Two stances, three genres, and four intractable dilemmas for the future of learning at scale. *Proceedings of the Seventh ACM Conference on Learning@ Scale*, 3–13.
- Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education, 26(2), 582–599.
- Schoonen, R., Snellings, P., Stevenson, M., & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. Writing in Foreign Language Contexts: Learning, Teaching, and Research, 77–101.
- Sevcikova, B. L. (2018). Online open-source writing aid as a pedagogical tool. *English* Language Teaching, 11(8), 126–142.

92

- Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, *27*(4), 657–677.
- Stanovich, K. E. (2009). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Journal of Education*, 189(1–2), 23–55.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296.
- Tan, B.-H. (2011). Innovating writing centers and online writing labs outside North America. *The Asian EFL Journal Quarterly*, *13*(2), 390–417.
- Tanil, C. T., & Yong, M. H. (2020). Mobile phones: The effect of its presence on learning and memory. *PloS One*, 15(8), e0219233.
- Wang, Y. (2019). Effects of L1/L2 captioned TV programs on students' vocabulary learning and comprehension. *CALICO Journal*, *36*, 204–224.
- Zhang, Z. V. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*, *43*, 100439.

Karmila Rafiqah M. Rafiq, Harwati Hashim, Melor Md Yunus

Faculty of Education, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia

karmilarafiqah@gmail.com, harwati@ukm.edu.my, melor@ukm.edu.my

Theoretical foundation in designing and developing a mobile app to support ESL learning for STEM learners

Biodata



Karmila Rafiqah M. Rafiq is a PhD TESL student at the Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She is a technology-enhanced language learning enthusiast. Her areas of interest are TESL, technology-enhanced language learning (TELL), and gamification in teaching English as a second language (ESL).



Harwati Hashim is a senior lecturer at the Centre for Innovation in Teaching and Learning, Faculty of Education, Universiti Kebangsaan Malaysia (UKM). She is an educational technology enthusiast and an m-learning practitioner. Her areas of concentration are TESL, mobile learning, mobile-assisted language learning (MALL), technology acceptance, innovative pedagogy, and the use of technology in teaching English as a second language (ESL).



Dr. Melor Md Yunus is a Professor and Deputy Dean of Research and Innovation at the Faculty of Education, Universiti Kebangsaan Malaysia. She is best known for establishing ICT integration in teaching and learning English as second language research. She is active in scholarly journal writing and publishing. The recipient of more than 200 innovation awards, she is highly regarded among her peers in her field. Her areas of expertise include English language education and innovative pedagogies.

Abstract

Computer-Assisted Language Learning (CALL) has become an emerging trend amid the pandemic alongside the rise of science, technology, engineering, and mathematics (STEM). However, the decline in STEM learners' pursuing STEM is worrying, and one reason is the decline in English language proficiency. Hence, this paper aims to review theories in design and develop ME4STEM (Mobile English for STEM) to enhance STEM learners' English vocabulary. This paper adheres to the Design and Development Research (DDR). The design and development of ME4STEM is underpinned by English for specific purposes, mastery learning, cognitivism, social constructivism, problem-based learning, and cognitive theory of multimedia learning theories. Referring closely to the theories, the development of ME4STEM consists of 6 elements: 1) learning videos, 2) practices, 3) reviews, 4) mini-assessments, 5) language exploration, and 6) discussions. The rationale for designing and developing ME4STEM is to encourage learners to learn is a personalized space and at their own pace, especially in this pandemic. ME4STEM and the

English language, leveraging opportunities for interdisciplinary study along with contributing to society by providing sustainable and novel English for STEM supplementary material. ME4STEM caters to all aspects of Smart CALL: personalization, contextualization, and socialization, thus contributing to emerging CALL trends.

Conference paper

Introduction

Computer-assisted language learning (CALL) has become an emerging trend amid the pandemic alongside the rise of science, technology, engineering, and mathematics (STEM). However, the decline in STEM learners pursuing STEM is worrying, and one reason is the decline in English language proficiency. This issue calls for a smart solution from the perspective of Smart CALL. The term Smart CALL includes personalization, contextualization, and socialization. To solve the language issue of STEM learners, this paper aims to review theories to design and develop ME4STEM (mobile English for STEM) to enhance STEM learners' English vocabulary. This study adheres to the Design and Development Research (DDR). Yet, before designing and developing ME4STEM, the theoretical foundation should be paid attention to, as a well-designed application should have solid theories.

Aspects of SMALL CALL

First, the personalization aspect of Smart CALL should be looked into. Personalization refers to adhering to learners of a specific profile. Due to that, ME4STEM is designed according to the needs of STEM learners, which caters to the personalization aspect. The needs of learners adhere to the Needs Analysis Model by Dudley-Evans and St John (1998). Findings from the needs analysis are fundamental in designing ME4STEM. The rationale for including the findings from the needs analysis is to ensure that the mobile application is personalized to STEM learners, enhancing language learning via a mobile app.

The second aspect of Smart CALL is the contextualization aspect. Considering this, ME4STEM should be contextualized, and learners could use the mobile app to learn vocabulary in their settings. ME4STEM is contextualized to STEM education. Using the cognitive constructivism theory (Piaget, 1974), social constructivism theory (Rieber & Wollock, 1997), and problem-based learning (Barrows, 1996), ME4STEM ensures that the content is authentic and contextualized.

Also, the design and development of ME4STEM includes the socialization aspect, which encourages learners to interact with peers and instructors via a forum in the mobile app. According to The Framework for the Rational Analysis of Mobile Education (FRAME), a mobile app's social aspect is essential. In ME4STEM, the socialization aspect is emphasized in the discussion section. Learners can freely use the forum platforms to chat, comment, and interact with peers to solve STEM-related problems using the language forms and functions. They are also encouraged to join a Telegram group for extra guidance and scaffolding by instructors. The rationale for including the socialization aspect is to ensure that the mobile app encourages active learning rather than unidirectional passive learning.

Theoretical foundation

The learning theories are chosen to provide a solid foundation for this study. Mastery learning is selected because it allows learners to review their learning. This theory underpins the mobile app's learning video, practice, and mini assessment. Cognitive constructivism is selected to ensure that learning modules are created with learners'

prior knowledge and background, focusing on assimilating previous experiences to new knowledge. This theory underpins the review element in the mobile app. The ESP functional language is selected because English for STEM requires authentic resources, thus rendering the importance of language for the specific context, and this theory underpins the language exploration element. Social constructivism is chosen as it allows interaction and discussion among peers to ensure the mastery of learning and is incorporated into the discussion forum. Since learners are discussing, there is a need to include problem-based situations to encourage learners to use the vocabulary in STEM-related contexts. This is because problem-based learning is constructivist (Barrows, 1996). These two theories underpinned the discussion element in the mobile app.

Positive feedback has been received from mastery learning, particularly in improving learners' various skills. The systematic method of presenting lessons, a step-by-step technique that helps learners properly comprehend the lesson, is one element of mastery learning (Desta et al., 2021). This is because the mastery learning strategy offers learners a thorough guide, improving their cognitive capacities to grasp the subject. In support of this claim, Komalawardhana et al. (2021) said mastery learning helped students learn the topics better. According to the findings of these studies, mastery learning offers students a chance to grow as individuals in the future. Orak and Al-Khresheh (2021) emphasize using technological advances to encourage English learners to accumulate knowledge by implementing cognitive constructivist strategies. It is also mentioned that optimal learning is achieved by learners applying their knowledge in real-world settings. Suhendi et al. (2021) support that cognitive constructivism can improve individual understanding by introducing a positive environment.

The materials used in ESP are authentic, have clear objectives, and are self-directed. One aspect of ESP is the authenticity of the contents. Dudley-Evans and St John (1998) said ESP learning materials do not have to be built from scratch. The authenticity of ESP materials is preserved since they are designed for a specific setting (Hutchinson & Waters, 1987). This demonstrates that ESP materials, created or utilized in language education, are unique in that they pertain to specific content for individual learners and settings.

Based on previous research, social constructivism allows learners to analyze their learning and make progress on their own (Amna Saleem et al., 2021). Du and Liang (2021) studied how social constructivism drives multimedia implementation in English classrooms. They say that the social constructivism of multimedia promotes English teaching and improves learners' learning ability. A study by Amna Saleem et al. (2021) mentioned that scaffolding could improve learners' critical thinking and encourage independent learning. However, the same study also showed that scaffolding could be built between learners. Learners who communicate with each other help create a meaningful learning environment. Tomak (2022) supports this claim by noting that learners who support peer learning tend to perform better and have higher critical and collaborative skills. In addition, these learners learn better from their peers who have better language skills than them because they feel more comfortable and can understand each other. This shows that scaffolding learning via interaction promotes a more desirable outcome.

In a study by Nurul Iskandar et al. (2021), English as a foreign language (EFL) for vocational students is practical through problem-based instruction. In addition to adequate language skills, vocational school students also need critical thinking and problem-solving skills. Because of this, they incorporated problem-based learning into the EFL classroom, providing learners with more opportunities to connect learning to their context. This is supported by Alemi et al. (2021), who reported similar results. In addition, research has shown that problem-based learning is fun (Jasti & Pavani, 2021) and develops students' understanding (Putra et al., 2021). Jasti and Pavani (2021)

conducted a study in an engineering school. They found that learners who understood problem-based instruction were more likely to enjoy problem-solving because they were allowed to structure their learning. A study by Putra et al. (2021) is consistent with this finding. They emphasize the importance of creating authentic problem-based tasks to ensure learners can think critically and solve problems based on their understanding of previous lessons. This highlights the importance of authenticity in lesson planning.

The rationale for choosing cognitive constructivism is to ensure that learning modules are created based on the learner's prior knowledge and context, emphasizing the ability to adapt the previous experience to new knowledge. This theory underpins the validation elements in this study. Effective study instruction is related to the learner's ability to discussion among peers to ensure mastery of learning and integration into discussion forums. As learners discuss, problem-based contexts need to be included to encourage learners to use vocabulary in STEM-relevant contexts. This is because problem-based learning is constructivist (Barrows 1996). The language form and functionality in the mobile app are based on purpose-built English and created so that the language can be applied to a specific STEM context (Dudley-Evans & St John, 1998). The rationale for creating a contextualized mobile app is to encourage learners to use language in familiar contexts, empowering English language learning.

Conclusion

Therefore, the design and development of ME4STEM caters to all elements in Smart CALL to contribute to emerging trends in CALL research. This research will guide researchers in designing and developing a language mobile application to cater to learners' current needs. Additionally, with a well-designed mobile application, everyone could access language learning, contributing to equal access to education as the Sustainable Development Goals (SDG) aspired.

References

- Alemi, M., Mohammadi, M., Ghanbar, H., Asghari, M., Rezanejad, A., & Tootoonchian, Z. (2021). EAP learners' attitudes toward problem-based learning: A fast-track to fourth generation universities. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, 9(2), 127–150. https://doi.org/10.22049/jalda.2021.27258.1313
- Saleem, A., Kausar, H., & Deeba, F. (2021). Social constructivism: A new Paradigm in teaching and learning environment. *Perennial Journal of History*, 2(2), 403–421. https://doi.org/10.52700/pjh.v2i2.86
- Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*, 1996(68), 3–12. https://doi.org/10.1002/tl.37219966804
- Desta, M. A., Workie, M. B. & Yemer, D. B. (2021). Investigating the effects of mastery learning instruction on medical students' academic writing skills proficiency. *Bulletin Monumental*, 22(8): 13–19. Retrieved from http://bulletinmonumental.com/
- Du, W., & Liang, X. (2021). The application of multimedia courseware in teaching college English intensive reading from the perspective of constructivism. *The International Journal of Electrical Engineering & Education*. https://doi.org/10.1177/0020720920984312
- Dudley-Evans, T., & St John, M. -J. (1998). *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge University Press.
- Hutchinson, T. & Waters, A. (1987). English for Specific Purposes: A Learning-Centred Approach. Handbook of Research in Second Language Teaching and Learning. Cambridge University Press. doi:10.4324/9781315716893

- Iskandar, N., Mustaji, Jannah, M., & Wicaksono, S. R. (2021). The problem based learning in enhancing students' critical thinking for reading skills in English teaching at vocational school. *International Journal of Recent Educational Research*, 2(2), 237–249. https://doi.org/10.46245/ijorer.v2i2.93
- Jasti, S. D., & Pavani, A. (2021). Employing problem based learning system in advancing communication skills proficiency in professional communication for engineering undergraduates. *Journal of Engineering Education Transformations*, 34, 2394–1707.
- Komalawardhana, N., Panjaburee, P. & Srisawasdi, N. (2021). A mobile game-based learning system with personalised conceptual level and mastery learning approach to promoting students' learning perceptions and achievements. *Int. J. Mobile Learning and Organisation*, 15(1): 29–49.
- Orak, S. D., & Al-Khresheh, M. H. (2021). In between 21st century skills and constructivism in ELT: Designing a model derived from a narrative literature review. *World Journal of English Language*, *11*(2), 166–176. https://doi.org/10.5430/wjel.v11n2p166
- Piaget, J. (1974). To understand is to invent: The future of education. Grossman.
- Putra, G. B. S., Padmadewi, N. N., & Santosa, M. H. (2021). Authentic materials used by senior high school English teachers. *Jurnal Imiah Pendidikan Dan Pembelajaran*, 6(1), 38–48. https://doi.org/10.23887/jipp.v6i1
- Suhendi, A., Purwarno, P., & Chairani, S. (2021). Constructivism-based teaching and learning in Indonesian education. *KnE Social Sciences*, 76–89. https://doi.org/10.18502/kss.v5i4.8668
- Tomak, B. (2022). Cultural and sociocultural elements of language learning for the Turkish learners of English in two different universities. *OPUS-Journal of Society Research*, 19(45), 116–128. https://doi.org/10.26466//opusjsr.1062860
- Rieber, R. W., & Wollock, J. (Eds). (1997). *The collected works of L. S. Vygotsky*. Springer. https://doi.org/10.1007/978-1-4615-5893-4

Mark R. Freiermuth, Ikuko Tomida

Gunma Prefectural Women's University, Tamamura-machi, Japan Khon Kaen University, Khon Kaen, Thailand

mark-f@fic.gpwu.ac.jp, tomida@kku.ac.th

Culture, confidence and connections: Telecollaboration as a springboard for successful JFL learning

Bio data



Mark R. Freiermuth is Professor of Applied Linguistics in the Department of International Communication at Gunma Prefectural Women's University in Japan. He is extremely interested in the psychology of L2 learners, especially regarding how learners are affected by the technologies they use to communicate with one another in the L2. His recent book and most of his other publications reflect these interests.



Ikuko Tomida is an adjunct lecturer in the Department of Eastern Languages at Khon Kaen University in Thailand. She has enjoyed teaching Japanese as a foreign language for many years. She is excited to offer collaborative opportunities to her Thai students.

Abstract

This study employed Zoom as a telecollaborative tool to connect 53 (28 on one day and 25 the next) Thai university students learning Japanese as a foreign language (JFL) with four Japanese university students. The study was an outgrowth of a concern about Thai students' waning confidence when using Japanese, and as such, we considered the activity employed in this study as an opportunity for them to develop more self-confidence. In addition, by sharing personalized cultural values with one another, it was hoped that their cultural acuity would be sharpened and enhanced via the interaction that took place. We also wanted to assess the value of using Zoom as a platform—especially regarding Breakout Rooms—a feature which allowed us to address the imbalance in student numbers from each side. The activity was a success regarding all three aims. In a posttest questionnaire Thai students indicated a boost in their L2 confidence, while both groups of students claimed they benefited from the mutual exchange of cultural ideas. In addition, Zoom worked quite well as a tool to facilitate the interaction. This is good news for L2 teachers, especially those in foreign language settings, who can use telecollaborative activities to increase their students' motivation and cultural learning.

Conference paper

Introduction

The development of cultural competence, as it relates to language learners, has become an area of great interest among L2 scholars. One of the problems, however, is how this
concept can be defined and framed in a way that maximizes the benefits to second language (L2) learners. To be sure, there is a strong attachment to the notion that to understand a language, one must also understand the culture from whence it is derived (Deardorff, 2006). Herein, there is a tacit assumption that culture and language come packaged together and are then delivered to learners. When considering the plight of foreign language learners in particular, there are three easily recognizable problems that are ineludibly adhered to this idea: 1) How can L2 learners internalize realistic cultural norms when they are removed from the source of the culture? 2) What exactly are the cultural norms that they are supposedly in need of developing? 3) Is culture simply a one-way street, flowing only from the target culture to the learners without any consideration of L2 learners' own cultural norms and values? Lacking a solid footing in the target culture, foreign language learners may not be that comfortable discussing concepts that they have only read about in textbooks or heard about from their teachers, which could have an adverse effect on their confidence when using the L2 (see especially Freiermuth & Huang, 2021). To address the concerns raised by the aforementioned questions along with their consequences, the present study investigates the effect of using a Zoom classroom, connecting Thai university students in Thailand who were learning Japanese as a foreign language (JFL), with Japanese university students in Japan.

Cultural Competences

As alluded to previously, although deemed essential to the effective learning of a second language, cultural competence is difficult to define. This is complicated by the fact that L2 learners cannot be accurately represented as a generic group with identical cultural profiles. Learners have different language learning experiences; they find themselves in different settings; they have different capabilities, and they have different backgrounds. In other words, what are considered standard cultural norms for a particular language must be balanced with what cultural notions the learners already carry with them into the L2 classroom. Culture viewed as a pot of tea with learners as the teacups being filled is a very narrow, unidirectional prism that ignores the learners' own complexities and distorts how real communication takes place. In other words, because successful conversation is founded on interaction, ideas are naturally exchanged and consequently, culture is also exchanged in the process. It was under these assumptions that Freiermuth and Huang (2021) conducted their research on culture. In their study, which connected Taiwanese and Japanese EFL learners together in dyads, culture was viewed as being something more personal to each individual rather than something that is taken from the pages of a textbook. Questionnaires given to the students confirmed that the learners developed a deeper intercultural understanding of their counterparts' cultural norms via the exchange of personalized cultural information. Students brought their own cultural knowledge with them and shared it with their peers. Freiermuth and Huang (2021) put it this way:

Stepping outside of existing cultural assumptions so that learners can view the world from different perspectives might mean making opportunities available for students to define cultural realities from their own perspectives—inside the classroom and even outside of it. The pot of gold at the end of the rainbow is that while students learn and practice their L2, they are simultaneously developing meaningful and more realistic intercultural competences alongside their language learning partners, which is a win for both teachers and their students. (p. 205)

To sum up, one of the fundamental aims of the present research was to offer a platform where students could experience an intercultural exchange with their overseas peers that considered the intricacies of individualized cultural expressions.

Motivation

From a purely pragmatic stance, a second aim of this study was to provide a comfortable learning space that would encourage L2 motivation. To pinpoint the problem, the Japanese foreign language teacher in Thailand had expressed concerns about her Thai students' waning confidence when using Japanese. There is strong evidence suggesting that perceived self-efficacy and confidence affect L2 learning. For example, Mills et al. (2007) found self efficacy of regulation to be the strongest predictor of success for the 303 intermediate learners of French in their study. In addition, in Piniel and Csizér's (2013) study, the researchers found that the 236 English language learners studying in a Hungarian secondary school were motivated by perceptions of their own abilities, which in turn were enhanced by their language learning experiences. As they astutely point out here:

Enhancing the learning experience by making the learning process more relevant and enjoyable for students will increase their self-efficacy, which will impact anxiety and intended behavior. Alternatively, anxiety reducing training helps students increase the amount of energy they invest into learning and that will make the process more enjoyable and their experiences more rewarding. Learning experience then affects students' self-efficacy beliefs, which in turn has both a direct and indirect impact through debilitating and facilitating anxiety on students' motivated learning behavior. (p. 539)

In other words, there was a cyclical effect—negatively perceived opportunities increased anxiety and debilitated self-efficacy as well as motivation, while positively perceived opportunities decreased anxiety and increased self-efficacy as well as motivation (see also Gardner et al., 1997).

Self-confidence is also associated with how learners view themselves. When considering Dörnyei and Kubanyiova's (2014) possible selves model, two components that point to sustaining motivation are relevant to the state of the L2 learners in this study. The first is the notion of strengthening learners' L2 visions. If L2 learners cannot easily identify a practical manner for learning the L2 by tying it to some kind of skill-building opportunity, their vision of their ideal L2 self may fade. A second, and closely related component of relevance, is the substantiation of learners' L2 visions. Because there are various hurdles to achieving language learning goals, individual experiences while using the L2 can lead a learner to conclude, "Yes, I can do this in the L2" or "No, I cannot do this in the L2"; the former psychological state tends to increase confidence and strengthen L2 vision while the latter obviously decreases and weakens it.

To sum up briefly, motivation was considered a key element in this study. Our intention was to provide an activity whereby the Thai JFL students could achieve L2 success while shoring up their L2 confidence and strengthening their language learning vision through their use of the L2.

Zoom

Using Zoom as a telecollaborative tool proved invaluable on several fronts. As a linguistic space for L2 learning, Zoom has shown itself to be relatively easy to use, and L2 teachers can easily record the interaction of students in the Zoom Lobby or in individual Breakout Rooms (with paid subscriptions all Breakout Rooms can be recorded). It has also been used to facilitate intercultural connections between dyads of EFL learners with much success. Freiermuth and Huang (2018, 2021) had Japanese and Taiwanese EFL learners resolve tasks together using the Zoom application, both of which turned out to be highly successful intercultural affairs.

An additional advantage of using Zoom involves Breakout Rooms. One nagging trouble spot for telecollaborative research in general is that classroom sizes may be quite large and are normally unbalanced, so if the aim is to promote some sort of student autonomy

Proceedings of the XXIst International CALL Research Conference

101

there are naturally complications (Do & Freiermuth, 2020; Freiermuth & Huang, 2021). In the present study, only four Japanese students were able to participate in the telecollaborative activity (on both days) compared to the 53 Thai students learning Japanese over two days (28 on the first day and 25 on the second day), so the Breakout Rooms were essential for group-like interaction to occur.

The rationale for using Zoom in this study then was threefold: 1) It has been a useful tool in the past for L2 language learning activities, 2) it has been successfully applied as a platform connecting different L2 learners together, and 3) it has Breakout Rooms which allow for semi-autonomous group work, which mitigates the problem of large class sizes and any imbalance in the number of participants between the collaborating parties.

Objective

The aims of this study then are as follows:

- 1. From participating in the L2 activity, would students gain a sense of cultural learning via the interaction with their overseas peers?
- 2. From participating in the L2 activity, would students become more confident and motivated to learn and use the L2?
- 3. From a purely technological angle, how would Zoom perform?

Methodology

Prior to the activity, both groups of students were given a questionnaire to assess their motivation level and perceptions of intercultural awareness.

The joint Zoom classes were scheduled to take place over two consecutive days at the same time each day. Of course, the time differences had to be considered and times were aligned in advance based on the Thai students' class schedules. 28 Thai second-year JFL university students took part in the study on the first day and 25 the second day. They were joined by four fourth-year Japanese university students who participated on both days. Four Zoom Breakout Rooms each housing six to seven Thai students and one Japanese student were used for the interaction phase.

Prior to the scheduled activity, all of the students (including the Japanese students) were asked to create and give a Powerpoint presentation using Japanese related to various aspects of what they considered to be culturally important to them personally. Both groups of students had practiced their presentations prior to the joint class. During the joint class, each participant gave their presentation in their assigned Breakout Room, which was immediately followed by a question and answer session. In most cases, due to time constraints, only two or three questions could be asked after each presentation. Each Japanese university student, performing the role of a 'tutor,' was asked to provide their own comments and questions during that time as a way to encourage the Thai presenter who had just finished. When all of the presenters had finished, students chatted among themselves in the Breakout Rooms in Japanese. Eventually, all of the students work place in just under two hours; however, on each day, a small group of students including the Japanese 'tutors' stayed to chat with one another in the Lobby for an extended period of time (depending on the student anywhere from about 5 to 30 minutes).

Following the activity, students were asked to complete an online posttest questionnaire (Survey Planet) on motivation and cultural awareness. The items were targeted to elicit answers aimed at addressing the aforementioned research questions using a 'before' and 'after' approach (i.e., the questions asked students to compare their perceptions prior to the activity and after it). The Results and Discussion that follow are intended to shine a light on the participants' answers.

Results and Discussion

In this short paper, we will look at the three primary elements as discussed in the literature to see if the international project had the desired effect. Hereafter, we will discuss the results obtained from the questionnaire data concerning cultural competences, motivation and innovation in the form of the Zoom application.

Cultural Exchange

Regarding cultural competences, 95.3% (41/43) of the Thai students mentioned that they enjoyed explaining Thai culture to their Japanese peers. In other words, they were very willing to explain these concepts in Japanese. Here are some of their comments:

- I think It's really great things to do that we can exchange how or what we think with other people from other country and can discuss other things too.
- ...this activity was awesome because I could talk with Japanese friends a lot. And I'm glad to teach you about Thai cultures with Japanese that was so challenged me.
- We talked about Thai culture while practicing Japanese language—listening, speaking, reading and writing. Talking with Japanese friends gave me peace of mind and more confidence in using Japanese. (translated from Thai)

As for the four Japanese students who participated, three of them said they enjoyed sharing their self-identified cultural norms with their peers. Here are the three positive comments:

- They were very interested in Japan and I was very happy. I also wanted to convey the goodness of Japan more. (translated from Japanese)
- I felt a little difference between people who were interested and people who were not, but I felt that most of the students listened intently and I was happy. (translated from Japanese)
- Because Thai students looks that they love Japanese culture. They heard my presentation earnestly.

These experiences align well with the conclusions drawn by Freiermuth and Huang (2021) who found that their Japanese and Taiwanese students enjoyed expressing personalized cultural norms with one another and were enriched by their experiences.

Motivation Levels

At the heart of this project was motivation-specifically the confidence-of the Thai students learning Japanese. The aims of this project revolved around the hope that such an activity would encourage the Thai Japanese language learners by boosting their confidence using Japanese to express something very familiar to them. The activity had the desired effect. 88.4% (38/43) of the Thai students indicated that their confidence levels were boosted while using Japanese with their Japanese peers. Only one student became slightly less confident and the remaining four indicated no confidence boost. That students were also asked about their Japanese confidence levels using the four macro-skills (i.e., reading, writing, listening and speaking), revealing very similar results. A related question asked the Thai students to assess their perceived anxiety levels regarding using the Japanese language in the future. 69.8% (32/43) of them predicted they would be less anxious about using Japanese during any future opportunities. Besides confidence and anxiety, we also wanted to know whether or not their overall motivation to study Japanese (39/43) was bolstered. 90.7% of students indicated that the activity had indeed increased their motivation to study Japanese. Here are some comments from students related their motivation:

- I feel that this project may make Thai students' confidence to use the Japanese language grow...(translated from Thai)
- This project is good for students. It developed many kind of our skills such as communication or critical thinking.
- This is a great project to exchange knowledge and culture, and strengthen our confidence. (translated from Thai)

Considering the Japanese students were playing the role of pseudo-advisors, they were asked about their motivation towards participating in this kind of project. All four students indicated that they were extremely motivated by the activity. Here are a couple of comments explaining their motivation boost:

- I could get much information about Thai culture. Thai students introduced many kind of cultures that I've never heard before. It's so interesting. I think Thai students too.
- I was surprised at the good skills of Thai students in Japanese, and I was told that they were enthusiastic about studying. I also wanted to improve my English. At the same time, I was motivated to learn not only English but also other languages. (translated from Japanese)
- I've exchanged contacts with some students, so I'd like to be able to communicate with my peers again sometime after the project. (translated from Japanese)

The results of this project are a testament to continuing this kind of project in the future. Many of the Thai students as well as the Japanese students mentioned their anxieties prior to the online meeting, but the interaction both within the Breakout Rooms and later in the lobby helped to vanquish their fears while simultaneously fueling their desire to improve their L2 skills. In other words, their L2 self-confidence and L2 self-esteem were both bolstered through such activities (Dörnyei & Ushioda, 2021). As Bandura (1997) has mentioned self-efficacy can be consciously boosted through any challenging success; in our case, the success achieved during the language learning experience cleared a pathway for further successes down the road.

Using Zoom Breakout Rooms

The final angle we examined was the innovation aspect. Zoom Breakout Rooms proved to be very easy to use and as a result, any worries about technological problems soon evaporated. Nonetheless, it must be understood that all of the students were quite accustomed to using Zoom. In fact, many of the students seemed quite comfortable employing various third-party 'mood-making' applications to make the Breakout Rooms more cheerful places to interact although these were applications that the researchers knew very little about.

The positive opinions about using Zoom for this kind of project were not limited to the researchers' opinions, however. The students also realized the benefits that Zoom provided. In all, 93% (40/43) of Thai students felt that Zoom was a good application for the learning of Japanese. They gave comments such as these:

- The connection was stable not only for videos and voices, screen sharing was good.
- Even if we are in different places, we can enjoy talking to each other. (translated from Thai)
- Zoom is simple and easy to use with many functions to choose from. (translated from Thai)

Japanese students agreed with their Thai peers, summed up in this one comment:

• Because it is easy to contact people.

104

To sum up, Zoom functioned well. Of course, some students had internet connectivity problems, but these incidents were few and far between. Nevertheless, even as technological advances are made, problems inevitably follow along (see especially Do & Freiermuth, 2020), but in many cases proper preparation can minimize the negative effects. That was our experience in this study and so, from the viewpoint of technological agency, Zoom served the students well.

Conclusion

In this short paper, we looked at three aspects of this telecollaborative project, namely cultural competences development, motivation and innovation. The project proved successful on all three fronts. Students enjoyed exchanging personalized cultural information; they were motivated by participating in the project and most importantly, Thai students' gained confidence to use Japanese; and finally, Zoom Breakout Rooms worked very well and made it easy to group students together despite the significant imbalance concerning numbers.

Of course there are some limitations that need to be mentioned. The study consisted of one two-hour period for the Thai students, so despite the positive sentiments they expressed, it is difficult to know the durability of their improved confidence and motivation. Longer term studies are always warranted and encouraged here by us as well. Also, this was done in an Asian context. Although there are many cultural differences between Japan and Thailand, there is also a palpable sense of camaraderie between the groups. To expand upon this research, we suggest using telecollaborative tools to connect L2 learners from a wide-variety of backgrounds encompassing various language learning circumstances.

As a final note, we are quite satisfied with virtually all aspects of this project. It is safe to say that there was trepidation on everyone's part as to whether or not this project could be 'pulled-off'. We can trumpet the successes from our own viewpoint but would rather give the students the floor. When asked whether or not they would like to continue this kind of project, all 47 (43 Thai and 4 Japanese) students hoped that we would continue this kind of interactive class, which can be summed up rather nicely by these two comments—one from each side: "Talking to Japanese people is a great opportunity to practice speaking Japanese and exchanging ideas. Therefore, I would hope this project could be organized again in the future" (translated from Thai) and, "We can learn how to talk and tell for people who do not know Japanese and Japanese culture much. It's very important learning, I think."

As a final pedagogical note, we would highly recommend this kind of project. From simple observation of the Breakout Room interactions, we could see that students were truly enjoying the chance to engage one another using the L2. In the end, it was very satisfying to observe L2 students relishing the opportunities to use the L2 in ways that were important to them, and that is really what is at the heart of language learning. Our recommendation to L2 teachers then is to open up every door that might lead students in that direction and this is what Zoom provided for our collaboration.

References

Bandura, A. (1997). Self-efficacy: The exercise of control. Freeman.

Deardorff, D. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, *10*, 241-266. https://doi.org/10.1177/1028315306287002

 Do, T. H., & Freiermuth, M. (2020). A need to communicate: An intercultural story of motivation generated in disrupted text-based electronic chat. In M. Freiermuth & N. Zarrinabadi (Eds.), *Technology and the psychology of second language learners and users* (pp. 467-490). Palgrave Macmillan/Springer Nature.

Proceedings of the XXIst International CALL Research Conference

105

Dörnyei, Z., & Kubanyiova, M. (2014). *Motivating learners, motivating teachers: Building vision in the language classroom.* Cambridge University Press.

Dörnyei, Z., & Ushioda, E. (2021). *Teaching and researching motivation*. Routledge.

- Freiermuth, M., & Huang, H. (2018). Assessing willingness to communicate for academically, culturally, and linguistically different language learners: Can English become a virtual Lingua Franca via electronic text-based chat? In B. Zou & M. Thomas (Eds.), Handbook of research on integrating technology into contemporary language learning and teaching (pp. 57-85). IGI Publishing. https://doi.org/10.4018/978-1-5225-5140-9
- Freiermuth, M., & Huang, H. (2021). Zooming across cultures: Can a telecollaborative video exchange between language learning partners further the development of intercultural competences? *Foreign Language Annals*, 54, 185-206. https://doi.org/10.1111/flan.12504
- Gardner, R., Tremblay, P., & Masgoret, A.-M. (1997). Towards a full model of second language learning: An empirical investigation. *Modern Language Journal*, 81, 344-362
- Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning*, 57, 417-442. https://doi.org/10.1111/j.1467-9922.2007.00421.x
- Piniel, K., & Csizér, K. (2013). L2 motivation, anxiety and self-efficacy: The interrelationship of individual variables in the secondary school context. *Studies* in Second Language Learning and Teaching, 3, 523-550.

Mei-Ya Liang

Department of English, National Central University, Chung-li, Tao-yuan, Taiwan

my_liang@yahoo.com.tw

Multimodal humor in emoji-mediated emotive communication

Bio data



Mei-Ya Liang is a professor in the English department of National Central University, Taiwan. Her recent research focuses on translingual socialization, multimodal storytelling, and computer-mediated discourse.

Abstract

Drawing on the features of sociality, mobility, and multimodality, the researcher presents emoji-mediated discussions among 43 university students in a Facebook group. The results of discourse-pragmatic analyses show the participants' frequent use of expressive and representative communicative acts and various types of multimodal humor, alongside laughing and thinking emojis. Utterance-final face emojis and other affective expressions have particularly been employed to make communicative moves and construct humorous discourses. This article discusses expressive and playful potentials of emojis in computer-mediated communication. This study may help to understand the impacts which social media are having on the students' development of digital literacy and affective linguistic practices.

Conference paper

Introduction

The widespread use of social media and the outbreaks of the COVID-19 pandemic have urged us to integrate digital literacy in language education while maintaining social relationships. Digital literacy involves textual knowledge, dispositions, and skills in multimodally and symbolically mediated practices, which provide new opportunities for developing creativity (Kern, 2021). The creative aspects of online text-based interactions are associated with the stylized performance to engage online users in learning to play in creative ways (Belz & Reinhardt, 2004; Danet et al., 1997; Liang, 2012; North, 2007; Warner, 2004). More recently, social media and mobile chat environments afford emojis and other multimodal signs, which allow online users to demonstrate verbal creativity and convey visual humor. Researchers have observed that digital artifacts, such as emoticons and emotis, co-occur with other verbal messages to perform social-emotive functions on various media and text-based computer-mediated communication (CMC) platforms, such as emails (e.g., Skovholt et al., 2014), weblogs (e.g., Kavanagh, 2016), Twitter (e.g., Spina, 2019), chats (e.g., Feldman et al., 2017) and Facebook (e.g., Konrad et al., 2020). Emojis have become conventionalized as illocutionary force indicators or tone markers to modify textual utterances (Dresner & Herring, 2010; Konrad et al., 2020). However, the playful aspects of digital literacy through the emotive

communicative acts of using languages and emojis as multimodal expressions have not been fully explored. To fully understand the playful potential of emoji-mediated communication, this study investigated university students' deployment of emojis in online discussions on news and media content, with special attention given to emotive communication and multimodal humor in a mixed-culture Facebook group.

Participants and Research Context

This study was part of a larger research project that aimed to explore online emoji-mediated communication among linguaculturally diverse university students as extracurricular activities (see Liang, 2022). This paper focuses on how emojis are used to mediate university students' humorous and playful discourses about various sociocultural issues and situations. Specific research questions are presented below:

- To what extent do the participants use face emojis in online discussions?
- How do the participants employ emojis to make emotive communicative moves?
- What types of conversational humor are produced in emoji-mediated interactions?

The research project recruited 43 university students (Appendix A), including three online moderators, 20 Taiwanese undergraduate students, and 20 international graduate students. The three moderators helped recruit participants from among their own acquaintances, friends, and friends' friends through snowball sampling. Accordingly, the relationship among the participants in this project could be viewed as a small social network of "weak interpersonal ties" (Milroy & Milroy, 1985, p. 364) through which we can observe the various innovative uses of a variety of emojis in online discussions.

The study examined emoji-mediated discussions in a private group created by the author. Although participants' interactions on the Facebook group are visible to one another, Facebook allows users to choose who can see their profile pages and posts and even to block other users so that they can keep certain personal information confidential. Discussions on Facebook can be viewed as a new form of sociality (Kern, 2021), which encourages user participation by friending, posting, liking, and tagging. Another two features-mobility and multimodality-are also essential because most Facebook users read text and watch images, photos, and videos through mobile devices, which allow constant connections and multimodal interactions (Tagg & Seargeant, 2016). Accordingly, participants in this study were asked to watch or read news and media content on Facebook and discuss them by posting their comments with emojis. During the eight-week discussions, two moderators took turns posting the discussion topics and questions (see Appendix B) by sharing multiple modes of interesting materials (e.g., texts, images, emojis, links, videos, and animations) on Tuesday mornings, and another moderator reminded the participants by tagging them on Friday mornings. The three moderators also socialized with the participants by liking and responding to their comments to establish and maintain social relations.

Data Collection and Analysis

The primary data were collected from online written discourses on Facebook for eight weeks over the course of the fall semester in 2020. The online discussions were first examined in terms of the number of words, comments, and emojis produced by the moderators, Taiwanese students, and international students each week. This revealed that the participants mostly used emojis at the end of utterances. The following analytic procedures were adopted. First, the analyses focused on the face emojis that occurred in the utterance-final position because they are closely associated with affect (Konrad et al., 2020) and, in particular, the humorous joking modality (cf. König, 2019). Second, after

recursive examinations of the data, the researcher and one research assistant coded the utterance-final emojis based on four communicative moves that the associated emotional expressions performed, i.e., (a) expressives: expressing what is inside the speaker's mind; (b) directives: directing interlocutors' actions; (c) representatives: representing what the world is like; and (d) commissives: committing to a future course of action. (cf. Scarantine, 2017). We first independently coded the data, examined the coding results, and achieved consensus on each coded item. Third, building upon previous studies on conversational humor and joking (e.g., Dynel, 2009; Langlotz & Locher, 2013; Messerli & Locker, 2021), the researcher identified types of conversational humor with utterance-final face emojis.

Results

The participants produced 4804 emojis over the eight-week online discussions. A total of 1486 comments were generated. The longest contained 232 comments (week 1), and the shortest contained 139 comments (week 7). The participants produced 71081 words. The three moderators produced 19947 words (M = 6649, SD = 2684.58) and 688 comments (M = 229, SD = 99.057), averaging 29 words per comment. The 20 Taiwanese students produced 21005 words (M = 1050.25, SD = 369.26) and 317 comments (M = 15.85, SD = 7.34), averaging 66 words per comment. Finally, the 20 international students produced 30129 words (M = 1506.45, SD = 962.37) and 481 comments (M = 24.05, SD = 17.13), averaging 63 words per comment. Compared to the other participants (n = 40, M = 19.95), the moderators (n = 3, M = 229.3) made statistically more comments (t = 3.65, p = 0.034) because they were required to interact with the participants by responding to their comments. Compared to Taiwanese, the international students (n = 20, M = 1506.45) produced statistically more words (t = 1.98, p = 0.030), which may be a strategy to make their ideas clear to unfamiliar discussion group members.

Use of emojis

Of the 1486 comments, 1290 (86.8%) contained emojis. Table 1 presents the 10 emojis most frequently used by the moderators, Taiwanese students, and international students. All of these emojis are face emojis. Various laughing emojis (e.g., rolling on the floor laughing and tears of joy) and the thinking face are found among the top 10 emojis across all three groups. The tears of joy emoji, which, coincidentally, was the Oxford Dictionary's Word of the Year in 2015 (see https://languages.oup.com/word-of-the-year/word-of-the-year-2015), was the most commonly used emoji among both the Taiwanese and the international students. Notably, the moderators used the rolling on the floor laughing emoji more frequently than the other two groups did, which may be evidence of a diffusion of change within the small weak-tie social network. The international students frequently used the beaming face with smiling eyes and various grinning emojis (e.g., the grinning face, grinning face with sweat, and grinning squinting face).

Mod	erate	ors				W1	W2	W3	W4	W5	W6	W7	W8	Total
1	Ø	rolling laughing	on	the	floor	38	52	21	26	10	2	49	40	238(5.0%)
2	8	Thinking				9	26	12	49	14	27	4	19	160(3.3%
3	2	Zany				9	4	7	0	16	0	11	3	, 50(1.0%)
4		Crying				6	6	3	4	15	4	0	5	43(0.9%)
5	::	Grinning				3	9	0	10	0	6	1	2	39(0.8%)

Table 1.	Тор 10	most freque	nt emojis
----------	--------	-------------	-----------

Proceedings of the XXIst International CALL Research Conference

7 winking 3 7 1 9 3 4 0 9 36(0.86) 8 winking 6 4 11 5 0 3 4 2 35(0.76) 9 wears of joy 3 7 1 5 1 0 14 3 34(0.79) 10 wopen mouth 5 4 3 6 0 3 0 6 27(0.69) Taiwanese Students w1 W2 W3 W4 W5 W6 W7 W8 Toi 1 was of joy 16 19 14 14 w3 110(2.3) 2 wolling on floor 14 8 10 14 14 14 14 14 3	%) %) %) <u>tal</u> 3%) }%
8 1 1 5 0 3 4 2 35(0.74) 9 2 tears of joy 3 7 1 5 1 0 14 3 34(0.74) 10 2 open mouth 5 4 3 6 0 3 0 6 27(0.69) Taiwanese Students W1 W2 W3 W4 W5 W6 W7 W8 Tot 1 2 tears of joy 16 19 14 16 8 4 20 35 132(2.8) 2 2 rolling on the floor 14 4 14 8 8 10 14 38 110(2.3)	%) %) <u>ktal</u> 3%) }%
9	%) %) <u>tal</u> 3%) }%)
10 • open mouth 5 4 3 6 0 3 0 6 27(0.6° Taiwanese Students W1 W2 W3 W4 W5 W6 W7 W8 Too 1 • tears of joy 16 19 14 16 8 4 20 35 132(2.8) 2 • colling on the floor 14 4 14 8 8 10 14 38 110(2.3)	%) <u>tal</u> 3%) 3%)
Taiwanese Students W1 W2 W3 W4 W5 W6 W7 W8 To 1 👄 tears of joy 16 19 14 16 8 4 20 35 132(2.8) 2 🎸 rolling on the floor 14 4 14 8 8 10 14 38 110(2.3)	o <u>tal</u> 3%) 3%)
1	3%) 3%)
2 🤣 rolling on the floor 14 4 14 8 8 10 14 38 110(2.3) 3%)
laughing	
3 🤗 Thinking 7 8 8 4 13 1 7 7 55(1.29	%)
4 😅 grinning face w/ sweat 6 3 4 3 1 1 6 3 27(0.69	%)
5 🙆 Monocle 0 5 8 1 5 2 0 3 24(0.5%	%)
6 😥 Crying 2 3 4 1 0 4 2 3 19(0.49	%)
7 😄 Zany 3 2 3 2 1 1 1 6 19(0.49	%)
8 🙀 exploding head 1 3 4 2 5 0 2 1 18(0.4	%)
9 😔 Pensive 3 3 4 1 3 3 0 0 17(0.49	%)
10 😆 grinning squinting 8 2 0 2 4 0 0 16(0.3°	%)
International Students W1 W2 W3 W4 W5 W6 W7 W8 To	tal
1 😂 tears of joy 17 9 22 27 18 12 24 31 160(3.3	3%
2 🤣 rolling on the floor 17 5 7 2 0 4 17 39 91(1.99) laughing) %)
3 ⇒ beaming face w/ smiling 12 12 10 14 8 8 6 8 78(1.69) eyes	%)
4 🤗 Thinking 6 5 9 18 3 1 6 15 63(1.39	%)
5 😅 grinning face w/ sweat 11 5 6 12 1 2 12 10 59(1.29	%)
6 😂 grinning squinting 11 1 4 13 1 4 4 3 41(0.99	%)
7 🙂 slightly smiling 8 3 11 5 7 4 2 0 40(0.89	%)
8 😳 smiling face w/ smiling 8 8 4 6 1 2 2 1 32(0.79	%)
9 😝 Grinning 8 5 3 5 1 1 3 3 29(0.6°	%)
10 😔 Relieved 8 1 2 5 3 1 4 3 27(0.69	

In line with Danesi's (2017) findings, we found that the international students tended to utilize face emojis with positive affects (e.g., laughing, smiling, grinning, and relieved). In addition to using laughing and smiling emojis, the Taiwanese students used a more varied set of emojis, such as face emojis with negative emotions (e.g., crying and loudly crying) and funny faces (e.g., zany and winking), which were not widely used among the international students.

As the objective of this study, participants were instructed to use emojis in the online discussions. For comparison purposes, the emojis used in the original Facebook comment threads in Week 3 and a set of online BBC news comment threads linked to in Week 5 are provided herein. Table 2 shows that the participants used a higher percentage of emojis than did general audiences on the original news and social media sites, which indicates the instructional effect of using emojis in this study.

Table 2. A comparison of emoji use

	Week 3 TikTok	Week 5 Mulan
	<u>comments with emojis/all</u>	<u>comments with emojis/all</u>
	<u>comments</u>	<u>comments</u>
Site data	152/3739 (30.81%)	113/1271 (8.89%)
Our data	141/160 (88.13%)	151/177 (85.88%)

Participants used face emojis, alongside verbal messages in their comments on a regular basis. In this study, which required the participants to use emojis, any lack of emojis in the interactions had certain implications. Some students may not have actively participated in the research study because they lacked pedagogical motivations. Other students did not use emojis when making serious responses or proposing conflicting ideas in discussions.

In what follows, the participants' online discussions with utterance-final face emojis will be analyzed to identify different types of emotive communicative moves and their functions.

Emoji-mediated emotive communicative moves

Table 3 presents the results of the four communicative moves used with the clause-final emojis in discussions over the 8-week study period. Overall, representatives (R) were the most common type (n=943, 40.84%), followed by expressives (E) (n=918, 39.76%), directives (D) (n=373, 16.15%), and commissives (C) (n=75, 3.25%). Here is one example:

W3-18CB: I just download Tiktok a week ago 😁 (R). It's pretty fun to use 😂 (E)...not just ban it because it's from a different country 😵 (D)... I plan to post some 😄 😄 😜 (C)

	W1	W2	W3	W4	W5	W6	W7	W8	Total
Moderators									
Expressives	50	74	17	45	44	28	32	41	331 (40.12%)
Representatives	36	23	34	33	27	16	24	23	216 (26.18%)
Directives	38	25	31	43	29	32	38	30	266 (32.24%)
Commissives	2	1	1	4	3	1	0	0	12 (1.45%)
Taiwanese Students									
Expressives	37	57	22	26	36	23	18	47	266 (42.36%)
Representatives	60	19	58	33	28	24	50	33	305 (48.57%)
Directives	7	6	7	3	7	3	1	2	36 (5.73%)
Commissives	5	2	4	2	4	2	2	0	21 (3.34%)
International Students									
Expressives	46	45	26	42	35	27	45	54	321 (37.50%)
Representatives	90	34	69	55	25	39	58	52	422 (49.30%)
Directives	9	10	8	13	13	5	4	9	71 (8.29%)
Commissives	15	4	2	8	3	6	2	2	42 (4.91%)

Table 3. Types of communicative moves with face emojis

In what follows, the participants' creative and playful potentials in expressing emotional or humorous stances will also be examined to see what types of humorous discourses are likely to occur in familiar and less close contexts.

Emoji-mediated humorous discourse

Proceedings of the XXIst International CALL Research Conference

After examining conversational humor in emoji-mediated communicative moves, 14 types of jokes and figures were identified (see Table 4). Common features of the humorous discourse included allusions (n=44), similes (n=42), irony (n=34), self-deprecation (n=27), witticisms (n=16), ingroup bond humor (n=10), hyperbolism (n=9), teasing (n=9), retorts (n=7), putdowns (n=6), wordplay (n=6), paradoxes (n=5), vulgarity (n=4), and metaphors (n=2). Allusion is the most common figure type employed by the participants in their discussions on news and social media content.

Terms (N)	Definitions	Examples
Allusion (44)	a citation alluding to a	W3-17-4CB: It's related to "Trump doesn't like
	pre-existing text or popular	it" 🥶
	culture artefact	
Simile (42)	an expression comparing two	W7-4JP: It's hard but amazing, like we are
	things using <i>like</i> or <i>as</i>	building the new Vavilon tower 🐸
Irony (34)	an expression that is the	W2-9-7WI: @AG I think the world knows what
	opposite of the literal meaning	China is doing, they just ignores it 🥴 wow,
		what a friendly world 😀
Self-deprecation	under-evaluating or criticizing	W2-5LE: a shame on me haven't seen it til
(27)	oneself	this day 🙂
Witticism (16)	a clever remark interwoven	W4-41-2JE @SN I think when the google
	into a conversation	finally get it, the restaurant is close. 😎 🤪
In-group humor	humor that enhances the	W1-18-2JU: @SN So cute! So we're both
(10)	in-group image or relationship	foodies! 😫 😳
Hyperbole (9)	an exaggerated statement that	W7-19 DA: Maybe we are so expressive so
	creates a strong impression	that all of our body try to communicate. 😁 😌
Teasing (9)	an utterance that makes fun of	W8-9-5KA: @RA the emoji you use seems
	others in a playful way	high 🤣
Retort (7)	a quick and witty response to	W5-2-4JO: It is not our rights that we are
	a preceding turn	defending, it is our sanity. 🍟
Putdown (6)	a remark to make something	W1-22-3 WI: Unfortunately, @VI gave back
	or someone foolish	the knowledge to her teacher
Wordplay (6)	making jokes by using words	W4-1-8VI: It's a good `die' `to die'!
	in a clever way	
Paradox (5)	a speech which contradicts	W6-22RA: I think it's pretty accurate overall
	itself	since I'm not a mathematician
Vulgarity (4)	utterances in vulgar quality	W7-19 DA: Some are dic* or a**hole too! 🤭
Metaphor (2)	an expression that implies	W6-20-8SN: @DA Well so the spring is still
	resemble between two things	very far away 🖲 🛛 🧃

Table 4.	Types	of jokes	and figures
----------	-------	----------	-------------

Discussions and Conclusion

This research study has drawn on the features of sociality, mobility, and multimodality to guide the emoji-mediated discussions in a Facebook group. While previous studies have explored the online participants' perceptions and intentions behind emoji use in online discussions through surveys or interviews (e.g., Liang, 2022; Konrad et al., 2020), this study focused on online discourses. As a result of the instructions on using emojis in intercultural communication beyond formalized learning situations, the university students frequently used various emojis to convey emotions and stances and to multimodally contextualize their virtual selves in the discussions of online news and social media content (for details, see Liang, 2022). The results of this study also showed that 86.8% of the online comments contained emojis and that representatives and expressives were commonly used with utterance-final emojis for exchanging worldviews and emotive stances in online socialization. Contributing to digital literacy practices that are broader than reading and writing, multimodal emoji text can be useful resources for developing embodied viewpoints on a range of sociocultural issues in creative ways.

This study found that the linguaculturally diverse university students produced a range of humorous jokes and figures through allusions, similes, and irony, along with various emojis, in discussions of socioculturally sensitive or worldwide moral or health issues (e.g., COVID-19). Specifically, the participants used various laughing emojis and the thinking face frequently. As shown in previous studies, emojis can function as humor-support indicators (Messert & Locher, 2021), which help in making communicative moves and establishing conversational humor in the dynamic discursive practices of online networked communication. Although humorous language play could be collaboratively developed and maintained by the interlocutors to negotiate relations between the interlocutors, the media, and the context (North, 2007; Tagg, 2013; Warner, 2004), it is argued that the discussion topics and prompting questions affect how the participants interact in emotive communication. In this study, some of the discussion topics were more humorous or emotionally charged than others (e.g., the discussions on English phrases commonly used in Taiwan that Americans do not understand in Week 8), and some topics were more politically sensitive for certain groups of students (e.g., the discussions on Taiwan's political status in Week 2 and the discussions the Hong Kong protests in Week 5). Accordingly, emojis were used more often with expressives to express affective stances than with the other three types of communicative moves.

Through social networks, the linguaculturally diverse participants can not only post texts and emojis in CMC but also play with linguistic and visual modes of expression on the move in their everyday lives beyond classroom learning. Emojis can be viewed as essential multimodal elements of emotive communication and humorous language play in online written discourses, which prompt the participants in the social situations to propose new perspectives on serious or controversial issues. This study has contributed to a broadened view of digital literacies by integrating the critical and creative multimodal practices into the study of university peers' online discussions that involves the use of emojis. This study may help to understand the impacts which social media are having on culturally diverse students' digital literacy and affective linguistic practices.

References

- Belz, J.A., & Reinhardt, J. (2004). Aspects of advanced foreign language proficiency: Internet-mediated German language play. *International Journal of Applied Linguistics*, 14(3), 324–362.
- Danesi, M. (2017). The semotics of emoji: The rise of visual language in the age of the internet. London: Bloomsbury.
- Danet, B., Ruedenberg-Wright, L., & Rosenbaum-Tamari, Y. (1997). Hmmm ... Where's that smoke coming from? Writing, play and performance on Internet relay chat. *Journal of Computer- Mediated Communication*, 2(4). https://doi.org/10.1111/j.1083-6101.1997.tb00195.x
- Dresner, E., & Herring, S. C. (2010). Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication Theory*, 20(3), 249-268.
- Dynel, M. (2009). Beyond a joke: Types of conversational humour. *Language and Linguistic Compass*, *3*(*5*), 1284-1299.
- Feldman, L. B., Aragon, C., Chen, N-C., & Kroll, J. (2018). Emoticons in informal text communication: A new window on bilingual alignment. *Bilingualism: Language* and Cognition, 21(1), 209-218.
- Kavanagh, B. (2016). Emoticons as a medium for channeling politeness within American and Japanese online blogging communities. *Language & Communication*, 48, 53–65.
- Kern, R. (2021). Twenty-five years of digital literacies in CALL. *Language Learning & Technology*, 25(3), 132–150.
- König, K. (2019). Stance taking with 'laugh' particles and emojis–Sequential and functional patterns of 'laughter' in a corpus of German WhatsApp chats. *Journal* of Pragmatics, 142, 156-170.

Proceedings of the XXIst International CALL Research Conference

- Konrad, A., Herring, S., & Choi, D. (2020). Sticker and emoji use in Facebook Messenger: Implications for graphicon change. *Journal of Computer-Mediated Communication*, 25, 217-235.
- Langlotz, A., & Locher, M.A. (2013). The role of emotions in relational work. *Journal of Pragmatics, 58*, 87-107.
- Liang, M. Y. (2012). Foreign ludicity in online role-playing games. *Computer Assisted Language Learning, 25*(5), 455-473.
- Liang, M. Y. (2022). Enregistering emojis in online socialization: Linguaculturally diverse university students' affective translingual practices. *Journal of Multilingual and Multicultural Development*. Advance online publication. https://doi.org/10.1080/01434632.2022.2036166
- Messerli, T. C., & Locher, M. A. (2021). Humour support and emotive stance in comments on Korean TV drama. *Journal of Pragmatics, 178,* 408-425.
- Milroy, J., & Milroy, L. (1985). Linguistic change, social network and speaker innovation. *Journal of Linguistics*, 21(2), 339–384.
- North, S. (2007). "The voices, the voices": Creativity in online conversation. *Applied Linguistics, 28*(4), 209–60.
- Skovholt, K., Grønning, A., & Kankaanranta, A. (2014). The Communicative functions of emoticons in workplace e-mails: :-). Journal of Computer-Mediated Communication, 19(4), 780-797.
- Spina, S. (2019). Role of emoticons as structural markers in Twitter interactions. *Discourse Processes*, 56(4), 345-362.
- Tagg, C. (2013). Scrapping the barrel with a shower of social misfits: Everyday creativity in text messaging. *Applied Linguistics*, *34*(4), 480-500.
- Tagg, C., & Seargeant, P. (2016). Facebook and the discursive construction of the social network. In A. Georgakopoulou & T. Spilioti (Eds.), Handbook of language and digital communication. NY: Routledge.
- Warner, C. N. (2004). It's just a game, right? Types of play in foreign language CMC. Language Learning and Technology, 8(2), 69-87.

Appendix A.	Participant	demographics
-------------	-------------	--------------

	Moderators	Taiwanese Ps	International Ps
	(n=3)	(n=20)	(n=20)
Gender			
Male	0	8	15
Female	3	12	5
Nationality			
Africa	0	0	1
America	0	0	3
Asia	3	20	14
Europe	0	0	2
Age			
20-22	3	20	0
23-31	0	0	20
Expertise			
Engineering	0	0	8
Science	0	0	10
Arts & social sciences	3	20	2
Level of Education			
BA/BS	3	20	0
MA/MS	0	0	13
PhD	0	0	7
Marital Status			
Single	3	20	19
Married	0	0	1
Length of Residence			
Less than 6 ms	0	0	1
6 ms-under 1 y	0	0	8
1 y-under 2 ys	0	0	5
2 ys-under 3 ys	0	0	6
3 yrs or more	3	20	0
Household			
Alone	2	7	11
Spouse/partner	0	4	1
Classmates	1	3	1
Friends	0	6	7
Online communication			
1-10 hours/week	1	7	9
11-20 hours/week	1	10	7
21-30 hours/week	1	3	6
Text/video chat			
1-10 hours/week	1	11	13
11-20 hours/week	2	2	2
21-30 hours/week	0	7	5
Social networking			
1-10 hours/week	1	7	8
11-20 hours/week	1	8	9
21-30 hours/week	1	5	3
21-30 hours/week	0	1	1

Appendix B. Discussion Topics and Questions

Week 1 2020/10/06 台灣婚禮(Taiwan Weddings) 黑素斯の熱吵店 (Taiwan Weddings Jesús's hot noisy shop) <u>https://youtu.be/h0UcsxxiZ8M</u> Have you ever heard of phatic words or cultural conventions in Taiwanese wedding banquets? How do you feel about the wedding etiquette? Share your family or friends' wedding stories with EMOJIS

Week 2 2020/10/13 A Quick Guide to Taiwan - TNNS <u>https://youtu.be/QaEBYegFABY</u> The host of the "Night Night Show," Brian, introduces Taiwan in an ironic tone. What figure(s) of speech do you like/dislike? Why ??

Week 3 2020/10/20 Trump says he will ban TikTok <u>https://www.facebook.com/5550296508/posts/10161147979141509/?extid=HRhd4MdqYMaHvxdg</u> <u>&d=n</u> Have you ever used apps which are operated by Chinese companies such as TikTok 抖音and Igiyi愛

Have you ever used apps which are operated by Chinese companies such as Tiktok 抖音and Iqiyi邊 奇藝? Do you think that the government should ban them? Why? 🤔

Week 4 2020/10/27 法式英文挑戰! (你聽得懂法國總統在說什麼嗎? feat. 路易 (French English Challenge! Can you understand what the French President is saying? feat. Louis) <u>https://youtu.be/t-pwpQMegMg</u> Have you observed different ways of speaking English when you watch online videos (e.g., speeches, TV dramas, films, games, etc)? Do you think accent matters in intercultural

communication? Share interesting stories with us 🐸

Week 5 2020/11/3 Liu Yifei: Mulan boycott urged after star backs HK police https://www.bbc.com/news/world-asia-china-49373276

What do you think about the hashtag <u>#BoycottMulan</u> on Twitter in response to Liu Yifei's pro-police comment on Chinese social media platform Weibo? Would you like to watch the film? Why?

Week 6 2020/11/10 Taiwan No. 1 for Covid response: Brookings study https://www.taiwannews.com.tw/en/news/4046909

Where is your county in this figure? Do you think it shows the current status of your country? Why ? If your country is not marked on the figure, show the current position of COVID-19 pandemic with a meme.

Week 7 2020/11/17

Why I moved back to Taiwan from Korea? 🔯 🏹 🔜 🍽 !? 😮

https://youtu.be/kGhtvEjrLEc

Have you been considered (im)polite when communicating with Taiwanese or people from other Asian countries? How do you mix English, other languages, or multiple ways of expression to

communicate with local people in the conversations? Share interesting stories with us ${f ige u}$

Week 8 2020/11/24

【我們看不懂英文!? 😜 】 美國人無法理解的台灣常用英文 | 彩曦&阿登

([We don't understand English!? $^{\circ}$] English phrases commonly used in Taiwan that Americans do not understand | Hailey & Adam)

https://www.youtube.com/watch?v=XniMHfCf0c4&app=desktop

What kinds of nonstandard Englishes have you ever seen or heard in Taiwan (e.g, slogans, ads,

public displays, etc)? Share interesting stories with us

Mengdi Wang, Ann Devitt, Juan Gao, Ciarán Baurer

Trinity College Dublin, Dublin, Ireland (Mengdi & Ann & Ciarán) Belvedere College SJ., Dublin, Ireland (Juan)

wangm3@tcd.ie, devittan@tcd.ie, gjuan@belvederecollege.ie, bauerci@tcd.ie

Computer mediated communication and task-based learning for adolescent learners of Chinese as a foreign Language in Ireland: An eBook task design under the adaption of Bridge 21 technology-mediated learning model

Bio data



Mengdi Wang is currently a PhD student in the School of Education at Trinity College Dublin. Her research interests are computer-mediated communications, telecollaborative language learning, technology-mediated task-based language teaching, Chinese as a foreign language learning and intercultural communicative competence.



Dr. Ann Devitt is an Associate Professor of Language Education at the School of Education in Trinity College Dublin. Her research interests lie in the area of language teaching and learning and technology enhanced learning. Dr. Devitt is currently academic director for Learnovate, the research and innovation centre focused on educational technology funded by Enterprise Ireland and hosted in TCD.



Ms. Juan Gao holds both a teaching certificate in Mandarin Chinese as well as a business management qualification background. Ms. Gao is an experienced and enthusiastic Chinese language and culture teacher and has joined Belvedere College S.J. since 2013. Currently, she lectures junior cycle Chinese and transition year Chinese courses in Belvedere College.

Mr. Ciarán Bauer moved into the educational sector and joined Bridge21, Trinity College Dublin, as Programme Manager in early 2012. Ciarán and his team offer a new model of learning that can be adapted for use in Irish secondary schools. His research interests are, 21st century learning, collaborative learning and the alignment of these methodologies with language learning methodologies.

Abstract

The Bridge 21 learning model emphasizes teamwork and technology mediation in the process of activity implementations. With the introduction of Chinese as a Leaving Certificate specification in the Irish secondary education system in 2020, there is a growing interest in Chinese language among schools, parents, as well as students. There are three types of Chinese language courses run through Irish secondary school settings: Junior Cycle (JC) Chinese Short Course, Transition Year (TY) Chinese and Leaving Certificate Mandarin Chinese. However, compared to other Anglosphere countries (e.g., UK, Australia), Ireland is in the early development stages of Chinese as a Foreign Language (CFL) learning (Osborne et al., 2019). Despite this, the open and flexible principle of Chinese course syllabi in JC and TY not only provides Chinese language

teachers with the freedom of selecting contents, but also makes it possible to adapt Bridge 21 model into teaching practice. However, Chinese as a curriculum specification at JC in Ireland has not been explored in depth or been combined with the Bridge 21 model. Therefore, this paper aims to elaborate on the design and implementation of an eBook activity which aligns to the Bridge 21 model in a JC Chinese course. The preliminary findings of participants' reflections suggest that the majority of participants had a positive experience in this activity and identified language development, especially recognition and production of Chinese characters, while one group of participants highlighted that they felt challenged working as a team. This may suggest there is a need for training of both technological tools as well as teamwork prior to conducting Bridge 21 learning activities in the future.

Conference paper

Introduction

Computer-mediated communication (CMC), as an important research strand of computer-assisted language learning (CALL), which refers to "communication that take place between human beings via the instrumentality of computers" (Herring, 1996, p.1). Researchers have investigated CMC since the 1980s (Thorne, 2007). Later on in 1990s, CMC was introduced and deployed in the field of second language acquisition and language settings (Beatty & Nunan, 2004; Herring, 1996; Li, 2018; Lin, 2015; O'Rourke & Stickler, 2017; Thorne, 2007; Warschauer, 1997). Generally, CMC is categorised into two different types, which are synchronous (SCMC) and asynchronous (ASCMC). Thorne (2007) characterised CMC as "multimodal, often (but not exclusively) Internet-mediated communication" (p. 1623). Multimodality as a key characteristic of CMC implies that an integration of multiple modes such as text, audio, and video embedded in specific CMC tools are individually or collectively utilised in language learning practice. The opportunities with more authentic target language exposure afforded by virtual environments and telecollaboration have been discussed among scholars and practitioners (Godwin-Jones, 2019; Lewis & O'Dowd, 2016; Milton, 2012; O'Rourke, 2007; Sadler, 2017; Schenker, 2012; Shih & Yang, 2008; Sykes et al., 2008). Therefore, these are identified as another two types of CMC having become prevalent after 2000.

The development of CMC in Chinese as a Foreign Language (CFL)

With the aim of investigating CMC tools integrated with tasks assisting Chinese as a foreign language learning, a systematic review was undertaken by authors who reviewed and analysed 69 research papers from 2008 to 2022. After screening and analysing selected papers, findings suggested that: 1) it is a common practice that collaborative learning facilitated by CMC tools among CFL learners mainly occurs in the same classroom environment (either physical or online), 2) a lack of studies targeting adolescent CFL learners addressing how CMC tools facilitate young learners of Chinese. Therefore, this paper will provide a valuable glimpse of the whole process of a collaborative eBook learning adolescent CFL learners in the Irish educational context.

Bridge 21 model

The Bridge 21 learning model is student-centred, which involves the following elements: technology-mediated, project-based, goal-oriented, social learning environment, learning space, and reflection. This model addresses a shift from teacher-centred to student-led instruction which aims to create a collaborative learning atmosphere for learners (Byrne et al., 2019; Lawlor et al., 2018; Sullivan et al., 2015). The full stages of the Bridge 21 activity model are shown in Figure 1. To date, the Bridge 21 model has been deployed in a wide range of post-primary school contexts in Ireland across a range of curriculum areas, including some European languages learning through Erasmus+ programmes

(Byrne et al., 2019). However, Chinese as a curriculum specification at post-primary schools in Ireland has not been investigated in depth or integrated with the Bridge 21 model.



Figure 1. The Bridge 21 Activity Model (excerpted from Bridge21 Introductory Handbook for Teachers, 2016, p. 13)

To address this gap, this paper describes an eBook activity as well as its implementation process in a secondary school located in Dublin. In addition, this paper will present preliminary findings about student participants' perceptions of the Bridge 21 model activity-eBook, which are based on individual and group reflection forms.

Background

This section will firstly introduce the background of CFL in the Irish educational system. Secondly, the profile of this eBook activity will be presented. Participants' demographic information will be available at the end of this section.

Chinese as a foreign language education in Ireland

Chinese as a foreign language was first introduced to Irish undergraduate students in 2006 (Wu, 2017). On the one hand, the development of Chinese in Irish higher educational institutes can be attributed to the support of two Confucius Institutes. On the other hand, some secondary schools themselves have expressed an interest in offering Chinese language courses to students in 2007 along with the establishment of the two Confucius Institutes in Ireland (Zhang, 2018).

Additionally, National Council for Curriculum and Assessment (NCCA) as a statutory body of the Department of Education in Ireland launched a taster Chinese course in 2012. This selective course mainly focuses on Chinese cultural aspects and a certain amount of linguistic content to Transition Year (TY) students who complete junior grades (i.e. Junior Cycle, hereafter JC) of secondary schools in Ireland. With the support of the Confucius

119

Institute, there were 262 classes in secondary schools and TY grades who run this course in 2017 (UCC Confucius Institute, n.d.) while UCD Confucius Institute (2016) reported that 102 schools ran a Chinese language course in 2016 (Osborne et al., 2019). In 2014, a pilot programme was tailored to JC adolescents. In 2017, Ireland launched *Languages Connect* - *Ireland's Strategy for Foreign Languages in Education 2017-2026*. This strategy aimed to introduce and develop Chinese as a curricular subject for Irish Senior Cycle students in the coming years. Following this national language strategy action, Chinese has been set up as a two-year curricular programme since 2020 (Curriculum Online, 2020). In June of 2022, students will take the first ever Leaving Certificate Chinese exam which is a state examination of the secondary school system and works as the university matriculation examination in Ireland. Table 1 provides an overview of Chinese language courses in secondary education in Ireland.

	Junior Cycle (3-year)	Senior Cycle (3 or 4-year) 15/16 Years Old Above					
	12-15 Years Old	Transition Unit (4 th year, optional)	Leaving Certificate (5 th & 6 th year)				
Title	Chinese Language and Culture	Chinese Culture and Language Studies	Leaving Certificate Mandarin Chinese				
Language and Culture Weight	70% language; 30% cultur	70% culture; 30% language	Not indicated but focus on accurately and fluently using Chinese appropriate to the age and stage of learning				
Study hours	100 hours	45 hours	180 hours				
Proficiency level (CEFR)	 A1.1 for speaking and listening Lower level for reading and writing 	Not indicated	Broadly aligned with Pre A1/A1 level of the CEFR				
Assessment	 Project/ a range of oral presentation styles, i.e. oral interview, role-play, poster presentation, digital presentation 	 Portfolio assessment Project work/presentation 	 An oral examination An aural examination A written examination 				

Table 1. An overview of Chinese language course in Irish secondary educational level

Due to the nature of school-based teaching and assessment for JC and TY, it is important to understand that Chinese language teachers have freedom to choose and design learning materials based on students' learning needs. The formats of assessment in these two courses are varied which allow Chinese language teachers and learners to possibly create a project-based teaching and/or learning environment. This relative laissez-faire attitude to Chinese learning paves the way for introducing the Bridge 21 model into teaching practice.

EBook activity

As mentioned above, project-based learning as an essential element of Bridge 21 model not only requires collaborative teamwork but also establishes an open and loose space for students to take control of the learning process, being accountable to themselves and other team members (Lawlor et al., 2018). In alignment with Bridge 21 model design

protocols, a four-week eBook activity was introduced and assigned to a group of student participants who studied a Chinese language and culture course in Ireland during the academic year 2021-2022. The CMC tool utilised in this eBook activity was Book Creator. Book Creator as a digital book making platform allows students to individually log on to the digital device and then start to collaboratively work seamlessly.

Participant information

There were 14 students (aged 14-15) voluntarily participating in this project. They are from a secondary school for boys in Dublin. This school uses iPads to assist students' daily learning. Participants have two years of Chinese learning experience in this course which is consistently lectured by a native Chinese teacher. The range of participants' Chinese language proficiency aligns with the Common European Framework of Reference (CEFR) and varies from Pre-A1 to B1 in terms of the four linguistic skills (speaking, listening, reading and writing).

EBook activity implementation

Considering the curriculum design and time restrictions, this eBook activity contains four stages of integrating the original seven-stage Bridge 21 model.

Stage 1: Set up and warm up

Participants firstly were randomly assigned into three groups with three to five members in the first week. A checklist of creating an eBook containing the main learning steps of the Bridge 21 model, which illustrates sub-tasks and outputs in each stage, was printed out and distributed to each group. Specifically, participants firstly were encouraged to name the team and assign specific roles (team leader, note taker, resource investigator, proof-reader) for the following activity. As this is a project-based learning activity, it is crucial for each member to understand their responsibility and take ownership of the learning artefact. The team leader is responsible for keeping the team on track and staying on task. The note taker contributes to making a record of brainstorming ideas from every member. The resource investigator seeks appropriate visual and audio elements as a content editor. The proof-reader will check the language input before the official eBook launch and presentation. It should be noted that every member can take on more than one role if they are comfortable.

In addition, each team decided page numbers as well as the eBook's title. To provide participants with an idea of previous linguistic knowledge and cultural content, a pool of topics were listed on the eBook checklist. Students discussed and confirmed the overall contents of the eBook by extending on given topics in the brainstorming process.

Students then were invited by the teacher to login to a closed-accessed digital library on Book Creator. A quick tutorial session was provided for students to navigate around Book Creator on their own. They became familiar with typing Chinese characters and Pinyin (Chinese pronunciation system), searching images, drawing, uploading photos as well as voice recordings. The arrangement of the tutorial session not only facilitated participants to familiarise themselves with Book Creator, but also paved the way for generating creative eBooks in the coming weeks.

Stage 2: Investigate and plan

At this stage, the whole team used the checklist received as a step-by-step handbook. Along with the determined topics in the first stage, students moved on to list the gist of the eBook contents. This step empowers each team to concentrate on constructing content and messages they would pass to the audience at the later presentation stage. Additionally, the following storyboard sketch phase allows the team to depict a rough version of the eBook ahead of contributing on Book Creator. According to feedback from participants, they demonstrated that this eBook checklist was handy and helpful to bring about creative ideas and guide them in making the eBook.

Stage 3: Create and Present

Following Stage 2, every member started to contribute to the allocated page by inserting language structures as well as illustrations. With the affordances of Book Creator, participants could use their own iPad device to seamlessly work on the same eBook without interruptions. In the final week of this activity, each team was invited to co-present their book in front of other classmates. To attract the audience's attention, two questions were given by the presenting group to challenge the rest of the classmates. This not only increased attention among other students while one group presented the book, but also created an active, engaging, and interactive environment.

Stage 4: Reflect

To grasp students' perceptions and personal experience of this activity under the adaption of the Bridge 21 model, the individual reflection form containing open-ended questions as well as a 5 Likert-scale questionnaire was administered to each participant after the presentations.

On the other hand, the group reflection focused on promoting team interactions through keeping a track of eBook activity achievements and challenges as well as attitudes toward using Book Creator. By the end, every group was required to assess teamwork and the eBook by following the eBook activity evaluation rubric. As this eBook activity is part of a PhD project, the focus group interview was conducted to invite participants to recall the whole process of eBook activity and freely talk about their opinions.

Reflections

This section will present students' perceptions of this eBook activity in terms of teamwork. The affordances and challenges will also be addressed.

Individual reflection on eBook experience

The analysis of item 'eBook activity overall experience' from individual reflection forms demonstrated a strongly positive experience among participants, with five rating it as excellent, five as good, and two as average. However, there were only two students who gave a fair rating regarding their own experience. This might be explained by the fact that one participant (Student 201) was a Chinese heritage speaker who found this activity unchallenging due to no emerging linguistic knowledge involved, while another student (Student 209) responded expressed low motivation and attitude towards the Chinese language course in his individual reflection form. Although those two students were not very positive, they still appreciated this opportunity to work with other classmates by rating 'strongly agree' to 'helping teammates when they needed' (Student 201), as well as 'listening to teammates' ideas' (Student 209) during the eBook activity.

Team Reflection on Achievements and Challenges

In terms of language learning and development, participants highlighted that this eBook activity helped them understand Chinese characters better by typing and recognising characters through iPad keyboard input during the process of making eBooks. Although the purpose of this activity was to review previous learning, participants demonstrated that they learned some new words and characters especially related to food and sports while preparing and writing pages on Book Creator.

As regards teamwork, one group pointed out that this was one of their major achievements. It is interesting that this group also recognised that the eBook activity provided them with an opportunity to "get to know a classmate better" as they introduced one member as a leading character in the book in front of the whole class. In terms of the role of Book Creator, the technology not only enabled participants to develop interpersonal relationships by communicating as a team, but it also brought the classroom together in a way that it had never been before. In addition, the adaption of Bridge 21 model in the eBook activity avails students' ability to research information as well as find reliable resources. Moreover, they reflected that the stage of investigation and planning was useful prior to designing a book on Book Creator.

Two challenges were reported in teams' reflection notes: language and teamwork. In terms of language barriers, one group addressed that "forming sentences of topics such as Irish foods and cultures is the problem as translations can be misinterpreted and translated incorrectly". This issue was identified by the authors in the presentation phase as 'hurling' was incorrectly translated to "投掷" in Chinese, while the correct answer would be "爱尔兰式曲棍球/板棍球". Another group noted that pronunciation was the main challenge. This might be explained by the fact that participants were not used to typing Chinese characters through Pinyin (Chinese alphabet pronunciation system). It was also seen within the eBook content as a certain amount of typos, e.g. tone differences: *和 (hé, and) instead of 喝 (hē, to drink), negative L1 interference: *吃床 (chīchuáng, eat bed) instead of起床 (qi chuáng, to get up), to name a few. On the other hand, this group noticed this because of the teacher's feedback on correct translations after the group presentations.

Regarding difficulties of teamwork, based on individual reflection form responses, although most participants enjoyed working with their team, one group highlighted working together was challenging. This may suggest that the training session on teamwork as well as team building warm-up activities is necessary. Another issue among participants was language barriers. L1 negative interference may bring spelling errors while participants type Chinese characters through Pinyin pronunciation regulations. In addition, the limited Chinese language knowledge may have hindered participants from contributing more creative and insightful content in this activity.

Conclusion

In the Irish post-primary educational system, Chinese has been introduced to adolescents through three courses at three schooling levels: Junior Cycle Short Course, Transition Year Course, and Leaving Certificate Curricular Subject. The flexibility in the Irish curriculum for Chinese allows Chinese language teachers to adapt innovative teaching and learning methods, such as the Bridge 21 model, into the classroom. The design and implementation of the eBook activity facilitated by Book Creator was tailored to a class of JC Chinese short course students in this study. Drawing on qualitative data from students, this project-based, collaborative eBook activity boosted participants' understanding of the language, in particular in relation to Chinese character writing. Furthermore, participants were very positive about the creative dimension of the eBook activity which allowed them to express themselves flexibly in Chinese. The structure of the eBook activity with its clear checklist works as a 'recipe' and scaffolds each group with indispensable 'ingredients' at different stages. The challenges in terms of teamwork arising in this activity would suggest the need to provide essential training of both technological tools as well as teamwork prior to conducting activities based on adapting the Bridge 21 model in future Chinese course activities.

References

Beatty, K., & Nunan, D. (2004). Computer-mediated collaborative learning. *System*, 32(2), 165–183. https://doi.org/10.1016/j.system.2003.11.006

Proceedings of the XXIst International CALL Research Conference

Trinity College Dublin. (2016). *Bridge21 introductory handbook for teachers*. http://tft-project.eu/index.php/handbook/

Byrne, J. R., Kearney, S., & Sullivan, K. (2019). Technology-mediated collaborative learning: The Bridge21 activity model in theory and practice. In L. Daniela (Ed.), *Didactics of smart pedagogy: Smart pedagogy for technology enhanced learning* (pp. 309–330). Springer. https://doi.org/10.1007/978-3-030-01551-0_16

Godwin-Jones, R. (2019). Telecollaboration as an approach to developing intercultural communication competence. *Language Learning & Technology*, 23(3), 8–28.

Herring, S. C. (1996). *Computer-mediated communication: Linguistic, social and cross-cultural perspectives.* Benjamins.

Lawlor, J., Conneely, C., Oldham, E., Marshall, K., & Tangney, B. (2018). Bridge21: Teamwork, technology and learning. A pragmatic model for effective twenty-first-century team-based learning. *Technology, Pedagogy and Education*, 27(2), 211–232. https://doi.org/10.1080/1475939X.2017.1405066

Lewis, T., & O'Dowd, R. (2016). Online intercultural exchange and foreign language learning: A systematic review. In T. Lewis, R. O'Dowd & S. L. Thorne (Eds.), Online intercultural exchange: Policy, pedagogy, practice (Vol. 1, pp. 21–66). Routledge.

https://elib.tcd.ie/login?url=http://search.ebscohost.com/login.aspx?direct=true &db=mzh&AN=2018230019

- Li, M. (2018). Computer-mediated collaborative writing in L2 contexts: An analysis of empirical research. *Computer Assisted Language Learning*, *31*(8), 882–904. https://doi.org/10.1080/09588221.2018.1465981
- Lin, H. (2015). Computer-mediated communication (CMC) in L2 oral proficiency development: A meta-analysis. *ReCALL*, 27(3), 261–287. http://doi.org/10.1017/S095834401400041X
- Curriculum Online. (2020). Mandarin Chinese. https://www.curriculumonline.ie/Senior-cycle/Senior-Cycle-Subjects/Mandarin-C hinese/
- Milton, J. (2012). Second language acquisition via second life. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (Vol. 10, pp. 1-9). https://doi.org/10.1002/9781405198431.wbeal1318

O'Rourke, B. (2007). Models of telecollaboration (1): eTandem. In R. O'Dowd (Ed.), *Online intercultural exchange* (pp. 41–61). Multilingual Matters. https://doi.org/10.21832/9781847690104-005

O'Rourke, B., & Stickler, U. (2017). Synchronous communication technologies for language learning: Promise and challenges in research and pedagogy. *Language Learning in Higher Education*, 7(1), 1–20. https://doi.org/10.1515/cercles-2017-0009

Osborne, C., Zhang, Q., & Xia, Y. (2019). *The past and present of Chinese language teaching in Ireland*, *2*, 23.

Sadler, R. (2017). Virtual worlds and language education. In S. L. Thorne & S. May (Eds.), *Language, education and technology* (pp. 375–388). Springer. https://doi.org/10.1007/978-3-319-02237-6_29

Schenker, T. (2012). Intercultural competence and cultural learning through telecollaboration. *CALICO Journal*, 29(3), 449–470.

Shih, Y. -C., & Yang, M. -T. (2008). A collaborative virtual environment for situated language learning using VEC3D. *Educational Technology & Society*, 11(1), 56–68.

Sullivan, K., Marshall, K., & Tangney, B. (2015). Learning circles: A collaborative technology-mediated peer-teaching workshop. *Journal of Information Technology Education: Innovations in Practice*, 14, 063–083. https://doi.org/10.28945/2176

Sykes, J. M., Oskoz, A., & Thorne, S. L. (2008). Web 2.0, synthetic immersive environments, and mobile resources for language education. *CALICO Journal*, 25(3), 528–546.

Thorne, S. (2007). Computer-mediated communication. In N. H. Hornberger (Ed.), Encyclopedia of language and education (pp. 1623–1634). Springer.

Proceedings of the XXIst International CALL Research Conference

- UCC Confucius Institute. (n.d.). *About us*. University College Cork. Retrieved 30 May 2022, from https://www.ucc.ie/en/asian/confucius/aboutus/
- UCD Confucius Institute. (2016). *The 11th Confucius Institute Conference Reference Materials (UCD CI)*. The 11th Confucius Institute Conference, Kunming, P.R. China.
- Warschauer, M. (1997). Computer-mediated collaborative learning: Theory and practice. *The Modern Language Journal*, *81*(4), 470–481. https://doi.org/10.1111/j.1540-4781.1997.tb05514.x
- Wu, Q. (2017). A survey report on the Chinese-teaching in college and university in county of Dublin in Ireland [Master thesis, Chongqing Normal University]. https://kns.cnki.net/KCMS/detail/detail.aspx?dbname=CMFD201801&filename= 1017168987.nh

Michael Hofmeyr

Osaka University, Osaka, Japan

mfhofmeyr@gmail.com

Digital game-based SLA in the wild: evidence from a qualitative case study

Bio data



Michael Hofmeyr teaches English for academic purposes and general linguistics at Osaka University, Japan. He has published research articles and book chapters on digital game-based language learning, innovative approaches for teaching critical thinking in an EFL context, and linguistic diversity in Japan. He is also pursuing a doctoral degree in applied linguistics at Kyoto University, looking at the potential of digital games to facilitate SLA.

Abstract

A substantial body of research in the field of digital game-based language learning (DGBLL) suggests that games may hold significant potential to facilitate second language acquisition (SLA). Since many CALL researchers are also language teachers, it is common for studies in this field to be carried out in a language classroom context, with a focus on implications for in-class learning and teaching. Scholars such as Sauro and Zourou (2019) have recently pointed out, however, that DGBLL is far more likely to take place outside of the formal educational context than within it. Of the billions of digital game players around the world, many play in languages that are not their L1 and this must surely contribute significantly to their language learning. To better understand how SLA may take place in the context of a cooperative multiplayer digital game, a case study was conducted in which four hours of spoken interactions between three learners playing a game face-to-face over several sessions were recorded, transcribed, and analysed. Other than stipulating the use of English, the researcher did not guide or structure the interaction in any way, so as to simulate play "in the wild". Interaction was analysed using two different qualitative approaches: a cognitive-interactionist analysis and a direct qualitative analysis. The first approach revealed occasional instances of negotiation for meaning and regular use of beneficial interactional strategies during interaction, while the second shed light on how unstructured game-based interaction among learners may facilitate the acquisition of L2 vocabulary and grammatical structures.

Conference paper

Background

A number of recent studies in digital game-based language learning (DGBLL) have emphasised the importance of supplemental classroom activities and materials in optimising learning outcomes (Peterson, 2021). In a classroom or language lab setting, a teacher may prepare worksheets to scaffold learning or to focus the attention of learners on language points relevant to the game-based task. Post-play debriefing sessions may also play a productive role in consolidating the acquisition of new language. Studies by Miller and Hegelheimer (2006), Ranalli (2008), and Wang (2019), among others, provide empirical support for the use of such activities to supplement game-based learning. In light of these findings, influential voices in the game-based CALL community (York et al., 2021; deHaan, 2022) have begun to advocate for a drastic shift in focus within the field, calling for a move away from research on the potential of games for *learning* and towards the pedagogy and praxis of game-based *teaching*.

It is certainly true that over recent years, many language educators and learners have come to accept digital games as a promising new tool in the teacher's pedagogical arsenal (Bolliger et al., 2015; Mifsud et al., 2013). However, it is also important not to forget that the vast majority of actual DGBLL probably occurs outside of the formal learning environment and thus well outside the control or influence of the language teacher, out in what Sauro and Zourou (2019) refer to as the "digital wilds". The reasoning behind this assumption is simple. Today, active players of digital games across the world are estimated to number well over two billion people (Narula, 2019). Since many popular games are not available in the first languages (L1s) of all players, and since many games require online language-based interaction between players situated across the world who speak a wide range of L1s, it follows that a large number of players, probably many millions, gain considerable exposure to languages other than their mother tongues via their gaming hobby. The engaging nature of digital games provides these players with a strong intrinsic motivation to reach a level of second language (L2) proficiency sufficient to successfully complete the challenges posed by the game and as such, digital gaming that takes place in informal settings is likely to contribute significantly to out-of-school language learning on a global scale. This hypothesis is supported by one survey study (Sundqvist, 2019) conducted in Sweden, which found that L2 English learners who regularly played commercial digital games as a pastime, as well as those who played specific genres of digital games, outperformed their non-gamer peers in various measures of English vocabulary. Apart from this study, however, research on DGBLL in informal learning contexts has been scant (Peterson, 2021).

In order to better understand the processes through which DGBLL occurs in the digital wilds, a case study was carried out to investigate the potential of one digital game belonging to a promising yet under-researched genre to facilitate second language acquisition (SLA).

Case Study: Keep Talking and Nobody Explodes

Game, Participants, Data, and Analyses

The game title used for this study was the commercially-produced cooperative puzzle game *Keep Talking and Nobody* Explodes (Steel Crate Games, 2015). The aim of the game is for a small group of players to work together to defuse a time bomb before the countdown timer on the bomb reaches zero. One player, the "defuser", has access to a screen displaying a bomb consisting of a combination of various puzzle modules. The other players, the "experts", have access to the *Bomb Defusal Manual*, which contains instructions on how to solve the game's various puzzles. The defuser may not look at the manual during play time, while the experts may not look at the bomb on screen. This creates an information-gap game mechanic, as players are required to quickly and effectively exchange verbal information in order to disarm the bomb before it detonates.

For this case study, three learner participants working towards a degree in English Studies at a large Japanese university played the game over four one-hour sessions. Two of the participants spoke Japanese as an L1 and the other spoke Mandarin, but this learner was also highly proficient in Japanese at the time. In order to simulate play in the wild and also to identify opportunities for SLA to take place during such game-based interaction, that is interaction not guided by a language teacher, the researcher aimed to

minimise interaction with and support given to the group of learners before, during, and after the play sessions. Learners started the first play session by completing a short in-game tutorial, but no further supporting materials or activities were provided. The participants were asked to use only English and the researcher did not intervene during play, except in cases where technical problems prevented participants from continuing. Audio and video recordings were made of the four entire play sessions, adding up to a total of four hours of recorded interactional data, which were then manually transcribed in full.

Two separate discourse analyses were performed on the data set to investigate whether and how the game facilitated SLA during the play sessions. The first analysis was informed by the cognitive-interactionist model of SLA and was performed in order to identify processes believed to facilitate language learning, such as the negotiation for meaning and associated interactional strategies. The second analysis followed an original analytical approach known as direct qualitative analysis. This approach was developed in order to identify probable acquisition episodes (PAEs) that could not be identified from the first analysis.

Analysis 1 Findings

In this section, the key findings of the cognitive-interactional analysis will be briefly described. A full description of these results can be found in Hofmeyr (2021).

The aim of the first discourse analysis is to investigate whether certain processes that have been hypothesised to facilitate SLA could be identified in the interaction data set. The frequent presence of such processes in learner interaction is taken as evidence that the game-based activity is likely to facilitate SLA. Negotiation for meaning occurs when communication between two or more participants in interaction, or interlocutors, breaks down, after which the interlocutors attempt to repair the breakdown by engaging in further dialogue (Long, 1996, 2015). In the data set, 51 negotiation episodes were identified. Episodes that were triggered by gaps in a learner's lexical knowledge or by vague language use were found to have elicited significant modified output and therefore to have held the most potential for SLA. While negotiation for meaning occurred several times during each one-hour play session, such episodes did not occur very frequently. The presence of negotiation for meaning in the game-based interaction is therefore not enough to justify the use of the game as an efficient means of language learning.

The analysis also identifies occurrences of interactional discourse management strategies, including confirmation checks, clarification requests, and comprehension checks (Ellis, 2008). Interactional strategies sometimes occur during negotiation episodes as attempts to repair a breakdown, or they may occur outside of negotiation episodes in order to improve clarity and avoid breakdowns from arising in the first place. Discourse analysis of the data revealed that confirmation checks, in which one interlocutor checks that they have understood another's utterance correctly, occurred very frequently during the game-based interaction, with more than a thousand instances of this strategy identified. However, the utterances produced as confirmation checks typically involved only the repetition of short phrases and therefore probably did not contribute much to the acquisition of new language. Clarification requests, on the other hand, prompted significant modified output and thus appears to hold greater promise for game-based SLA. However, a total of only 35 such instances were identified in the data set, indicating that the game offers only limited potential to benefit SLA in this way. Comprehension checks occurred only four times over the course of the four play sessions and in none of these cases prompted rich linguistic output, leading to the conclusion that their effect on SLA was negligible during play.

Analysis 2 Findings

In the second analysis, six PAEs were identified in the data set and investigated in detail, three involving the acquisition of L2 vocabulary and three more involving the acquisition of L2 morphosyntactic features. In order to identify interactions beneficial to SLA that did not involve negotiation for meaning or the use of interactional strategies, a new approach to discourse analysis was devised. This approach is referred to as direct qualitative assessment and it attempts to minimise reliance on speculative SLA theory by identifying self-evident instances of SLA during learner interaction. This is done by pinpointing interactions in which one interlocutor clearly lacks knowledge of a given L2 feature, after which another interlocutor correctly models the feature in question, after which the first interlocutor reproduces the feature fully or partially. The following example from the data set illustrates what a PAE might look like in practice. While working to solve a maze puzzle during the first play session, one learner produced the following utterance to indicate the position of a red triangle appearing on an on-screen six-by-six-square grid: "From top one, two, three, four [...] and from right, two". A more appropriate syntactic structure to describe the location of the red triangle would be: "Fourth from the top and second from the right". During the second play session, when the learners encounter the same type of puzzle again, a different learner indicates the position of the red triangle with the utterance "Second from the right and third from the bottom", employing the appropriate syntactic structure. The learner who was at first unable to produce this structure repeats the appropriate structure two times after the model provided by his peer. Later during the second play session, this learner again encounters a maze puzzle and this time, he produces the appropriate structure spontaneously, without a peer first modelling the correct form. Tracking the learner's output in this way reveals that over the course of two play sessions, he had progressed from not being able to produce the appropriate structure at first, to then repeating it accurately after an interlocutor, and finally to producing the appropriate form spontaneously. This result demonstrates that even in cases where neither negotiation for meaning nor conventional interactional strategies occur in game-based learner interaction, opportunities exist for rich linguistic exchanges that hold the potential to facilitate SLA.

Hofmeyr (in press) provides a more complete description of the direct qualitative approach to discourse analysis and explains the reasons for its development. Detailed analyses of the six PAEs from the data set are also provided. In addition to improving our understanding of how cooperative puzzle games based on an information-gap mechanic may help learners to acquire and consolidate L2 vocabulary and grammatical structures, the analysis also produced findings to suggest that game-based interaction may improve L2 pronunciation. While the phonological analysis carried out for this study was not sufficiently robust for publication, the potential of game-based spoken interaction to improve L2 pronunciation is a topic ripe for further study.

Conclusion

The findings of the two analyses described here strongly suggest that face-to-face spoken interaction between learners playing an interactive puzzle game without the support of a language teacher or supplementary materials can facilitate SLA in several ways. The game used in the case study elicited negotiation for meaning as well as interactional strategies that sometimes led to the production of modified L2 output. However, it is still uncertain whether these potentially beneficial processes occurred frequently enough to efficiently produce positive learning outcomes. The game-based interaction also enabled learners to notice and to address gaps in their knowledge of vocabulary, morphosyntactic features, and possibly also pronunciation features of the L2. Finally, it should be noted that the conditions under which data was collected for this study did not constitute a perfect simulation of gaming in the wild, as the play sessions took place in a specially reserved classroom on a campus, the researcher was present during the play sessions, and learners were instructed to use only English while playing

the game. In a genuinely informal setting, learners playing a cooperative or other type of multiplayer game would be free to communicate in any language they choose and any given learner would be likely to choose to do so in their second or foreign language only if they did not share a first language with the other players. In online play involving players of different nationalities or cultural backgrounds, communication is likely to take place in English or another regional lingua franca. It therefore follows that DGBLL opportunities in the wild would be most readily available to players whose L1 were not the lingua franca. In the case of face-to-face cooperative games such as *Keep Talking and Nobody Explodes*, friends playing the game together would be more likely to share the same L1 than players from different locations being randomly matched up to play with an international pool of partners or competitors online, as is often he case with popular commercial multiplayer games. However, in self-contained multicultural settings such as a university campus that houses a sizable proportion of international students, learners might play the game in a shared L2 and thus most likely also benefit from the processes of SLA demonstrated in this case study.

This paper calls for a wider awareness among the CALL and DGBLL research community of the impact of digital gaming in the wild on foreign language development on a global scale as well as a recognition that in-class DGBLL must in all likelihood only account for a small fraction of the language learning that happens by means of digital games. This is not to deny, however, the positive potential of in-class DGBLL and that of supplemental activities designed to enhance the learning of specific target L2 structures. Han and Reinhardt (2022) reiterate this point in claiming that investigating DGBLL in the wild can "shed light on the potential role of language educators in leveraging activity in the digital wilds for formal learning purposes". Further research on this topic can help us to better understand what kinds of digital games and gameplay mechanics best facilitate SLA and how to exploit and enhance these games and mechanics for the purposes of in-class learning. It can also help to clarify when it would be most advantageous to integrate digital game-based learning activities into the classroom and when it would instead serve learners best to encourage DGBLL outside of class, whether at home with their friends or in a quasi-formal learning environment such as a self-access learning centre or a language laboratory.

References

- Bolliger, D. U., Mills, D., White, J., & Kohyama, M. (2015). Japanese students' perceptions of digital game use for English-language learning in higher education. *Journal of Educational Computing Research*, *53*(3), 384–408. https://doi.org/10.1177/0735633115600806
- deHaan, J. (2022, June 17–19). *Games and play in (researching) language teaching and learning: What really matters?* [Keynote address]. JALTCALL 2022 Conference, Japan.
- Ellis, R. (2008). *The study of second language acquisition* (2nd ed.). Oxford University Press.
- Han, Y., & Reinhardt, J. (2022). Autonomy in the digital wilds: Agency, competence, and self-efficacy in the development of L2 digital identities. *TESOL Quarterly*. Advance online publication. https://doi.org/10.1002/tesg.3142
- Hofmeyr, M. (2021). Lighting the fuse for interaction and negotiation: The potential of information-gap digital puzzle games for language learning. *Technology in Language Teaching & Learning*, 3(1), 1–20. https://doi.org/10.29140/tltl.v3n1.450
- Hofmeyr, M. (in press). Direct qualitative assessment of language learning outcomes in digital game-based interaction: An exploratory case study. In M. Peterson, & N. Jabbari (Eds.), *Digital games in language learning: Case studies and applications*. Routledge.

- Long, M. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie, & T. K. Bhatia (Eds.), *Handbook of research on language acquisition* (Vol. 2, pp. 413–468). Academic.
- Long, M. (2015). Second language acquisition and task-based language teaching. John Wiley and Sons.
- Mifsud, C. L., Vella, R., & Camilleri, L. (2013). Attitudes towards and effects of the use of video games in classroom learning with specific reference to literacy attainment. *Research in Education*, 90(1), pp. 32–52. https://doi.org/10.7227/RIE.90.1.3
- Miller, M., & Hegelheimer, V. (2006). The SIMs meets ESL incorporating authentic computer simulation games into the language classroom. *Interactive Technology* and Smart Education, 3(4), 311–328. https://doi.org/10.1108/17415650680000070
- Narula, H. (2019, December 29). A billion new players are set to transform the gaming industry. *Wired*. Retrieved from

https://www.wired.co.uk/article/worldwide-gamers-billion-players Peterson, M. (2021). Digital simulation games in CALL: A research review. *Computer Assisted Language Learning*. Advance online publication. https://doi.org/10.1080/09588221.2021.1954954

- Ranalli, J. (2008). Learning English with The Sims: Exploiting authentic computer simulation games for L2 learning. *Computer Assisted Language Learning*, 21(5), 441–455. https://doi.org/10.1080/09588220802447859
- Sauro, S., & Zourou, K. (2019). What are the digital wilds? Language Learning & Technology, 23(1), 1–7. https://doi.org/10125/44666
- Steel Crate Games. (2015). Keep Talking and Nobody Explodes [Computer software]. https://keeptalkinggame.com

Sundqvist, P. (2019). Commercial-off-the-shelf games in the digital wild and L2 learner vocabulary. Language Learning & Technology, 23(1), 87–113. https://doi.org/10125/44674

Wang, Q. (2019). Classroom intervention for integrating simulation games into language classrooms: An exploratory study with the SIMS 4. *CALL-EJ*, 20(2), 101–127.
York, J., Poole, F. J., & deHaan, J. W. (2021). Playing a new game—An argument for a teacher-focused field around games and play in language education. *Foreign Language Annals*, 54(4), 1164–1188. https://doi.org/10.1111/flan.12585

Myung Jeong Ha

Sangmyung University, Cheonan, Republic of Korea

mjha@smu.ac.kr

Using Socrative to facilitate problem-based learning in large undergraduate courses

Bio data

Myung Jeong Ha is an associate professor of English Language and Regional Studies at the College of Global Humanities at Sangmyung University of Korea. Her research interests include computer-enhanced language learning, computer-mediated communication, and corpus-driven discourse analysis.

Abstract

This study investigates students' perspectives on a student response system (SRS), Socrative, in a sizeable university-level course. The purpose of utilizing Socrative was to enhance participation and promote collaboration among students when they interact with one another to discuss their knowledge and critical thinking of the content of the course. Participants of the study were 82 students enrolled in an American Culture and Society course at a university in Korea. Web-based questionnaires and classroom observations were employed to examine students' perceptions of Socrative use and their engagement in the related PBL activities. Overall, the students' evaluation of the effectiveness of Socrative in a large enrollment class was more positive than negative. The students believed that the use of Socrative increased their interests and engagement, improved critical thinking skills and conceptual understanding of the course contents, promoted interactivity and participation, and allowed them to get motivated. The statistical analysis results showed little gender difference for all the items except for interactivity.

Conference paper

SRS Technology-Supported PBL

Although student response systems using clickers have been around since the 1960s, they have only recently been given attention as tools to promote learning, mainly focusing on the active learning approach. Grounded in active learning, constructive pedagogy claims that students learn more successfully when they are expected to actively build their understanding of course concepts (Anderson, 1987). It is thus teachers' responsibility that creates learning environments where students can practice applying and discussing course concepts during class hours.

Moreover, SRS is viewed as mobile technology that enhances problem-based learning (PBL) in extensive courses (Hoekstra, 2008). PBL has historical origins in medical education but has been used in various discipline-related academic studies, including architecture, business, engineering, law, and science in universities (Savery, 2015). PBL is defined as an instructional approach intended to facilitate prior knowledge activation, critical analysis of arguments, and promote a deep understanding of the scientific perspective (Hmelo-Silver, 2004; Loyens et al., Kirschner, & Paas, 2012). PBL is based on learning the principles of constructivism and emphasizes the learner's active participation

in the learning process (Savery & Duffy, 1995). The PBL pedagogy shares some standard features with those highlighted by utilizing SRSs in the classroom.

Unlike traditional lecture-style classes, this approach motivates students to learn through involvement in a real problem. Research and theory in psychology suggest that by having students learn through the experience of solving problems, they can acquire content and thinking skills (Hmelo-Silver, 2004). Notably, the PBL approach stresses that social interaction is essential to knowledge construction, acquisition, and application (Evensen & Hmelo, 2000). The social negotiation of meaning lies at the core of the knowledge construction process. PBL proponents suggest that instructional designers create learning environments where the teacher provides guidance and support, and the learner's knowledge construction is facilitated (Hmelo-Silver, 2004; Tseng et al., 2012).

Thus, the effectiveness of content learning can be maximized when learners are actively engaged in social interaction, such as group activities and interpersonal communication. Hoekstra (2008) emphasizes that the PBL approach stimulates active student involvement during the learning process by placing students into small groups where they work to apply course concepts. Numerous studies have demonstrated that group discussions encourage students to explore specific topics, process material more deeply, and create meaning in the material (Kirschner et al., 2009; Prince, 2004).

At the same time, the very aspects of PBL that allow for a productive learning experience make it more challenging to be implemented in a large classroom. The common concerns that discourage instructors from implementing at the undergraduate level include the nontraditional teacher role, the atypical student role, and potentially challenging group interactions (Aarnio et al., 2014). Students may feel uncomfortable when transitioning from passive roles in the traditional-lecture classroom to the leaders of their self-directed learning experiences. Students may also have trouble working in groups if their prior academic experiences were individual and not collaborative. These concerns over PBL may be highly alleviated by relying on SRS technology, Socrative, which has been shown to create a comfortable environment that allows all students to participate anonymously (Benson et al., 2016; Stowell et al., 2010).

Informed by previous research, this study utilized SRS technology to allow PBL to be implemented as a supplement to regular, didactic coursework. In this vein, this study contributes by investigating students' perceptions of the use of Socrative in the PBL enacted classroom environment. By doing so, this study attempts to make a significant bridge between SRS technology and the PBL approach. Very few studies attempted to explore SRS technology connected with PBL enactment in higher education contexts. The significance of the present study is that it expands on the scope of some notable work carried out previously that has focused on SRS technology.

Methods

This study is part of a larger study that investigated an alternative pedagogical approach to large classes using Socrative. Socrative-mediated PBL was implemented in the American Culture and Modern Society class offered by the Department of Language and Literature. The undergraduate content course was an elective introductory course of which the target audience was not limited to the Department of Language and Literature. Students typically did not have accounts to use online systems; when the teacher was logged into the system, they entered the teacher's online classroom with her code. When students provide answers on Socrative through their smartphones, the answers are instantly uploaded to the teacher's screen on PC. The screen is shared through the overhead projector to the whole class as the students engage in the activity so everyone can check the entire class's progress. Relying on the PBL approach, this class includes a small group discussion which is followed by a topic-related movie watching.

Proceedings of the XXIst International CALL Research Conference

Socrative was used as a formative assessment tool to review content through pre-made quizzes. Students were requested to choose the preferred answers by using Socrative. Before dealing with the main topic for the day, review quizzes were given first. During the next 10 minutes, a background check of the main topic for the day was provided. Subsequently, the actual lecture for the day was presented, which was followed by the related movie watching. After the movie watching, topic-related issues were given for student discussions. Students were given five minutes to think over the issues individually and participated in group discussions for 15 minutes. Students were expected to post the common results of discussions on a Socrative platform. Each group's responses were shared in Excel on the data projector screen for a whole class discussion. The class discussion was allotted 15 minutes, and during the discussion, equal participation was encouraged by the instructor. The instructor spent the remaining time providing feedback and finished the lecture by summarizing the essential points.

Data Collection

A mixed-method design was used, combining a background information survey and end-of-semester questionnaires, with classroom observation. At the beginning of the semester, a web-based background survey was administered. The primary purpose of the background information survey was to find out the students' level of education, their major, gender, learning style preferences, and their previous experiences with any other course having an SRS component. Regarding learning style preferences, Reid (1998)'s self-reporting questionnaire was used with modifications. According to Reid's classification, the students who show a group learning style learn more easily when they study with others and complete learning tasks through group interactions. In contrast, the students who have individual learning preferences learn best when they work alone and make better progress in self-study. These two learning styles show apparent differences in the learning process— group interaction vs. self-study. The use of Socrative in this study is highly associated with group discussions, and these two learning styles were selected as crucial individual difference variables. At the end of the semester, the online survey created in Google Docs was distributed to quantitatively investigate the students' perspectives on the classroom activities conducted through Socrative and preferences in using Socrative. The survey questions were drawn and appropriately adapted from previous studies (Cardoso, 2011; Dervan, 2014; Guarascio et al., 2017).

Data Analysis

The Statistical Package for the Social Science (SPSS) 17 was used to carry out a t-test and Analysis of Variance (ANOVA). The alpha level for all statistical analyses was set at 0.05. To compare the mean scores between the female and the male group, an independent sample t-test was performed. An independent sample t-test was also used to measure differences between the group of the individual learning style and the group of group learning styles concerning the items of interests, critical thinking, engagement, conceptual understanding, interactivity, participation, and motivation. A one-way ANOVA was carried out to analyze mean differences among three disciplines regarding the use of Socrative in the large class. Finally, open-ended responses were analyzed qualitatively by the researcher.

Results

First, altering classroom dynamics by enhancing students' participation and discussions with Socrative use plays certain roles in male and female students' learning. Regarding a research question of gender difference in the use of Socrative, both male and female students felt positive about the items of interests, critical thinking, engagement, interactivity, conceptual understanding, active participation, and motivation. The p-values from the t-test provided little evidence of the gender gap associated with using SRS

technology in terms of interests, critical thinking, engagement, and conceptual understanding. Interestingly as for interactivity between peers and a teacher, there was a significant difference between female and male students. Whereas the PPT lectures systematically presented information about American culture in a neutral and impersonal manner, the short questions from Socrative situated the cultural knowledge of American society in a specific personalized context.

Second, statistical analysis showed little difference between students of individual learning styles and group learning styles in terms of interests, critical thinking, engagement, conceptual understanding, and interactivity. As for the items of participation and learner motivation, in contrast, the mean of student perception was higher in students of group learning style (M=4.33, M=4.0) and significantly different between groups (F=1.473, p<.05; F=2.055, p<.05). This implies that the use of Socrative goes better with students who show a preference for group learning to keep them motivated and increase the degree of participation. It is, therefore, reasonable to raise the instructor's awareness of the different learning styles of class members in advance.

Third, the one-way ANOVA results showed no significant differences in the mean values of three different disciplines at .05 level. Overall, the students from three different disciplines showed positive responses. When asked whether they would recommend further use of Socrative in future classes, they also responded positively.

Finally, the main strengths that the students mentioned are related to shared opinions and thoughts and thus co-constructed knowledge grounded in intersubjectivity. The findings show that the students considered the ability to share their opinion with the whole class to be the strongest aspect of using Socrative in a big classroom. For most students, using Socrative made the learning environment feel more cooperative in a lecture class. Through Socrative-mediated group discussions, students helped each other by evaluating each other's reasoning and catching each other's way of thought. As for the weaknesses, most of them were associated with technical difficulties, not classroom usage. With students being so accustomed to online applications, there were few comments in a survey related to software usage difficulties.

Implications

Based on the findings, the following suggestions are made for using Socrative in a large content-based class. First, in using Socrative, teachers should provide students opportunities to explore and internalize learning content and related issues by their views and thoughts. At the same time, it is essential to keep in mind that students are interested in sharing peer responses. Second, educators and practitioners must utilize appropriate pedagogical approaches incorporated with Socrative use. A well-designed pedagogical strategy, including a sense of one's learning goals and how to achieve these goals utilizing Socrative help, is of great importance. Third, using Socrative brings up new problems and challenges to consider—that is, how to deal with student resistance to increased learner accountability and connect Socrative activities with student grades. Accordingly, it is a crucial responsibility of the instructor to make sure that students keep on the right track of Socrative-mediated activities.

References

Aarnio, M., Lindblom-Ylänne, S., Nieminen, J., & Pyörälä, E. (2014). How do tutors intervene when conflicts on knowledge arise in tutorial groups?. Advances in Health Sciences Education, 19(3), 329-345.

Anderson, C.W. (1987). Strategic teaching in science. In B.F. Jones, A.S. Palincsar, D.S. Ogle & E.G., Carr (Eds.), *Strategic teaching and learning: Cognitive instruction in*

Proceedings of the XXIst International CALL Research Conference

135
the content areas (pp. 73–91). Alexandria, VA: Association for Supervision and Curriculum Development.

- Benson, J. D., Szucs, K. A., & Taylor, M. (2016). Student response systems and learning: Perceptions of the student. Occupational therapy in health care, 30(4), 406-414.
- Cardoso, W. (2011). Learning a foreign language with a learner response system: The students' perspective. *Computer Assisted Language Learning*, *24*(5), 393-417.
- Dervan, P. (2014). Increasing in-class student engagement using Socrative (an online Student Response System). *AISHE-J: The All Ireland Journal of Teaching and Learning in Higher Education, 6*(3).
- Evensen, D., & Hmelo, C. E. (Eds.). (2000). Problem-based learning: A research perspective on learning interactions. NJ: Erlbaum, Mahwah.
- Guarascio, A. J., Nemecek, B. D., & Zimmerman, D. E. (2017). Evaluation of students' perceptions of the Socrative application versus a traditional student response system and its impact on classroom engagement. *Currents in Pharmacy Teaching and Learning*, 9(5), 808-812.
- Hmelo-Silver, C. E. (2004). Problem-based learning: What and how do students learn? Educational Psychology Review, 16(3), 235-266.
- Hoekstra, A. (2008). Vibrant student voices: Exploring effects of the use of clickers in large college courses. *Learning, Media and Technology, 33*(4), 329-341.
- Kirschner, F., Paas, F., & Kirschner, P. A. (2009). Individual and group-based learning from complex cognitive tasks: Effects on retention and transfer efficiency. *Computers in Human Behavior*, 25, 306–314.
- Loyens, S. M. M., Kirschner, P., & Paas, F. (2012). Problem-based learning. In S. Graham, A. Bus, S. Major, & L. Swanson (Eds.), *Application to learning and teaching: Vol. 3. APA educational psychology handbook* (pp. 403-425). Washington, DC: American Psychological Association.
- Prince, M. (2004). Does active learning work? A review of the research. *Journal of Engineering Education, 93*, 223–232.
- Reid, J. M. (1998). Understanding learning styles in the second language classroom. Upper Saddle River, NJ: Prentice-Hall Regents.
- Savery, J. R. (2015). Overview of problem-based learning: Definitions and distinctions. In A. Walker, H. Keary, C. E. Hmelo-Silver, & P. A. Ertmer. (Eds.), *Essential readings in problem-based learning: Exploring and extending the legacy of Howard S. Barrows* (pp.5-16). Purdue University Press.
- Savery, J. R., & Duffy, T. M. (1995). Problem-based learning: An instructional model and its constructivist framework. *Educational technology*, 35(5), 31-38.
- Stowell, J. R., Oldham, T., & Bennett, D. (2010). Using student response systems ("clickers") to combat conformity and shyness. *Teaching of Psychology*, 37, 135–140.
- Tseng, K.-H., Chang, C.-C., & Lou, S.-J. (2012). The process, dialogues, and attitudes of vocational engineering high school students in a web problem-based learning (WPBL) system. *Interactive Learning Environments*, 20(6), 547-562.

Orit Ezra, Anat Cohen, Dror Shany

Tel Aviv University, Tel Aviv, Israel Tel Aviv University, Tel Aviv, Israel Tel-Hai College, Qiryat Shemona, Israel

oezra1@mail.com, anatco@tauex.tau.ac.il, drorsha@m.telhai.ac.il

A mobile diary application as an instrument for collecting real-world and real-life contextualized language learning

Bio data



Orit Ezra is a Ph.D. student at the Department of Education in Mathematics, Science and Technology, School of Education at Tel Aviv University, where she also completed her Master's degree. Her research focuses on mobile technologies for second language learning, particularly Chinese. She had lived and studied Chinese in Taiwan to further improve her proficiency.



Anat Cohen (PhD) is a senior academic staff member of the School of Education at Tel Aviv University, Head of the Technology and Learning program in The Department of Education in Mathematics, Science and Technology, and Deputy Chair holder of technology and education affairs in the UNESCO Chair in Technology, Internationalization and Education (TIE).



Dror Shany (PhD) is a certified educational psychologist and lecturer at the department of special education, Tel Hai College, Israel. His expertise includes school psychology, L1 early literacy, and learning disabilities. He also has a special interest in adult L2 language learning, especially of the non-phonetic Chinese and Japanese languages.

Abstract

This study is part of a long-term project dealing with the definitions, measurements, factors, and outcomes of outside-classroom real-world (RW) and real-life (RL) contextualized (mobile-assisted) language learning, that is, CLL and CMALL. A known research obstacle in the field is how to collect learners' RW/RL contextualized learning given 'noisy' data. There is an issue of memory for self-reported data, which weakens questionnaires' validity. The experience sampling method (ESM), however good in assisting learners' recall, is guite laborious. Diaries are introspective, and asking about recent events, such as a day before, as in the day reconstruction method (DRM), can reduce memory bias. Furthermore, digital tools facilitate data analysis for researchers. Therefore, in this study, we developed a diary application to collect data on learners' experiences in outside-classroom RW-RL CLL & CMALL. We developed the diary app based on the DRM guideline and eight principles that make it usable in future studies regardless of the target language with no need for coding. The main learner functionality (reporting events) is based on the RW/RL contextualized learning model adjusted to the diary. The diarists are guided unequivocally based on the model and using clear instructions and questions about aspects of language, context, materials, and feelings. The Admin Monitor facilitates data tracking and analysis. Thus, this pioneering language-learning diary application for outside-classroom RW-RL CLL and CMALL is more structurally consistent and comprehensive than traditional language-learning diaries. The app is purposeful for researchers and practitioners interested in using language-learning diaries or learning about students' outside-class RW-RL CLL & CMALL.

Conference paper

Background

Outside-class RW-RL CLL & CMALL

A thread of studies has developed over the years, dealing with contextualized language learning (CLL) that involves learning *outside of class* in one's daily real-world (RW) and real-life (RL) (e.g., Edge et al., 2011; Hyland, 2004; Litzler, 2014b). Studies exploring such out-of-class RW-RL-CLL have also been given attention in mobile-assisted language learning (MALL), given mobile devices' utility in contextualized language learning (Comas-Quinn et al., 2009; Kukulska-Hulme, 2012; Pegrum, 2014; Sharples et al., 2007). Thus in recent years, a vast collection of studies examining such real-world and real-life contextualized mobile-assisted language learning (RW-RL-CMALL) has appeared (e.g., C. M. Chen & Li, 2010; Ezra & Cohen, 2018; W. Y. Hwang et al., 2014; Ibrahim et al., 2017; Lee, 2019; Lee & Park, 2019; Rivers, 2009; Sandberg et al., 2011; Santos et al., 2016; Zhang, 2018).

Data collection instruments of outside-class RW-RL CLL & CMALL

Outside class, RW-RL CLL & CMALL can occur at any place or time during the day. Therefore, learners' reports on their daily language learning occurrences capture this type of learning. Studies have used various data collection instruments for learners' reports, including Time-Tracker web tools (Viberg et al., 2021), questionnaires, surveys, and interviews (Bradley et al., 2017; Cohen & Ezra, 2018; Lamb & Arisandy, 2020; Steel & Levy, 2013).

Studies adopting the experience sampling method (ESM) (Edge et al., 2011; Larson & Csikszentmihalyi, 2014) can overcome memory limitations and an inherent bias due to self-report methods (Bradburn et al., 1987). Diaries and journals are also considered valuable for recording occasions of language learning beyond the classroom while maintaining immediate recollection by recording them shortly after the event (Reinders & Benson, 2017). Moreover, introspective journal entries provide data that may reveal aspects of the language learning experience which are challenging to attain by other means (Curtis & Bailey, 2009). Several language learning diary studies in second or foreign language learning environments looked into learners' outside-classroom activities (Benson et al., 2018; Chanjavanakul, 2017; X.-B. Chen, 2013; García-Amaya, 2017; Hyland, 2004; Krishnan & Hoon, 2002; Litzler, 2014a, 2014b; Litzler & Bakieva, 2017b; Ranta & Meckelborg, 2013; Warden et al., 1995). These diary studies illustrate how learners recalled outside activities focusing on *language, context, materials,* and *feelings* and what purposes keeping a diary can serve.

However, the literature to this date, to the best of our knowledge, has provided no instrument to systematically collect learners' reports on their outside-classroom RW-RL CLL & CMALL. Thus, an instrument built on consistent guidelines and clear definitions of RW and RL learning, with aspects of language, context, materials, and feelings, is missing. Assessing and evaluating learning of this kind of activity is indeed a challenge (Comas-Quinn et al., 2009).

The language diary app development

Objectives

Given diaries' merits, such as evoking recent memories, and in light of their potential superiority over ESM in terms of lower participant burden (Karapanos, 2020), we chose to implement our outside-class RW-RL CLL & CMALL instrument as a diary. Our diary only

accepts events that happened the same day, or one day before the learners record the data, as suggested in the day reconstruction method (DRM) (Kahneman et al., 2004). Furthermore, in light of the need for structured diaries (Brown et al., 2015) to overcome memory and definition difficulties (Vavoula, 2005), we drew on the RW/RL contextualized learning model (Cohen & Ezra, 2018; Ezra & Cohen, 2018). We thus guide the learners through questions, as is customary in both ESM and DRM. We collect all types of outside-learning events, with or without a mobile device. We elicit information about aspects of language, context, materials, and feelings to fortify recall and form a more holistic picture of one's contextualized learning. Finally, we are developing our diary as a mobile application for two main reasons. First, electronic, written journal entries significantly facilitate data analysis (Curtis & Bailey, 2009). Second, mobile devices' noticeable characteristics, including permanency, accessibility, immediacy, and interactivity (Viberg & Grönlund, 2012), facilitate an ideal ongoing basis for filling in diaries (Stockwell, 2022).

Thus, the aim of this study was to 1. develop a holistic, structurally consistent data-collection diary instrument, 2. develop a self-explanatory, handy diary application, and 3. develop a diary application that empowers researchers and practitioners to track, analyze, and easily maintain required changes and configurations. We thus designed and developed the "Langy" (**Lang**uageDiary) mobile application.

Design principles and adjusted RW/RL contextualized learning model

Eight crucial *design principles* were recognized for the diary app in light of the study objectives. Namely, we strived to

- 1. Boost memory
- 2. Provide a structure-based diary
- 3. Offer a self-explanatory diary
- 4. Build a holistic diary
- 5. Use common LL diary-studies tips
- 6. Facilitate data tracking
- 7. Facilitate data analysis
- 8. Support system configuration

A detailed explanation of the eight design principles follows. We boost learners' memory by adapting the DRM method and using some other mnemonic tools (1). Structured input consistency is delivered using the RW/RL contextualized learning model (see below in detail), guiding questions, and annotations enriched with examples from our previous studies. These can also boost memory and help overcome definition discrepancies among learners. Furthermore, the RW/RL contextualized learning model (Cohen & Ezra, 2018; Ezra & Cohen, 2018) was adjusted to fit the structured LL diary. According to the adjusted model, the diary provides learners with a consistent platform for recording RW and RL information about their learning events (2). A handy diary app with unequivocal guidance is provided using guiding guestions and annotations, including ample examples (3). All aspects - language, context, materials, and feelings - are included in the diary. A holistic picture might further fortify learners' memory (4). We comply with diary privacy important to learners' candidness, and provide language scaffolds such as explicit instructions to ignore grammar and spelling mistakes (Curtis & Bailey, 2009). We also strive to push learners toward ongoing filling in diary records rather than postponing and filling multiple records collectively (Stockwell, 2022) (5). The app admin user can monitor data recording of learners (6) and analyze their data (7). The diary can be modified and configured according to desired characteristics of the diary study.

The diary app modules

The following section describes how the principles, the RW/RL model and the DRM one-day-backward guideline, were introduced into the two modules of the diary, namely the *Learner Diary* and the *Admin Monitor*. The UI components presented in the screenshots below were configured for our first diary study on learning Chinese as a second language.

The Learner Diary module - The Opening screen

Figure 1 shows the Opening screen in the Learner Diary module. The user pseudonym is presented in line with the privacy considerations of the instrument. Researchers and practitioners can modify the title 'Chinese learning diary' to fit other language diaries. In the following screens, any place indicating Chinese is modifiable to meet the studied language. A message informing the diarist how many days remain before the allotted time ends boosts the app's self-explanatory nature. The calculation is based on a pre-configured number of days in which researchers or practitioners want to hold their program. Diarists are guided to create a mobile event (MALL event) or non-mobile event (LL event). Configurable annotations, placed on both buttons (with or without mobile), highlight the definition of a language learning event, the differences between these two event types, and provide some typical examples. They serve as a consistent structure to follow and a mnemonic device to trigger learners' recall of relevant events. Configurable annotations are also helpful in increasing the app's self-explanatory capability and are attached to many fields in the app. The app instructs learners from diverse countries in simple English and asks learners to ignore their spelling and grammar mistakes while writing, adhering to the language scaffolds principle. The learner can open a maximum of four events in parallel (a configurable number). It serves the principle of ongoing filling in records by encouraging learners to create an event shortly after it happened. On the other hand, learners are pushed to fill in until completion by limiting the number of opened events and shutting records at the end of the day.



Figure 1. Learner Diary module - Opening screen

The "Create MALL/LL Event" flows

According to our initial configuration, the questions' MALL and LL event flows are similar. The questions concerning mobile device information and mobile learning materials are not included in the LL event flow. Researchers and practitioners can design and configure their own desired flows and respective questions and immediately deploy the app's changes without requiring any app update. Figure 2 shows the first screen in both event types – General Event Details, in which the learner selects the event date as the date of the report or one day earlier and records the place where the event took place and a short event name.

In addition to its memory-aid functionality, the DRM's one-day-backward restriction is also helpful in promoting the app's self-explanatory nature by focusing students on the relevant dates. A configurable annotation, adjacent to the place text field, explains to the learner the places relevant to the study with examples of excluded places (such as in school or at home as in our first case study). It also instructs the learner to mainly record English and use the target language (Chinese in our first study) to elaborate. To bolster the learners' memory, they are instructed to record more precisely by specifying location type and name, and some examples are provided. This instruction facilitates the data analysis as researchers and practitioners can differentiate between events that happened on the same day in similar places. The short event name can trigger learners' memory. The configured event name annotations instruct the diarist to enter a short name in English; in the other text fields, diarists are instructed to enter information mainly in English with the target language (e.g., Chinese) to elaborate when required, to avoid diary-caused CMALL or CLL. Learners are instructed to record more precisely by specifying the language activity (e.g., shopping for groceries or eating in a restaurant), and typical examples are provided. Thus, the app's self-explanatory and memory requirements are bolstered. The next button instructing learners to proceed likewise supports the app's self-explanatory principle.



Figure 2. Learner Diary module - General Eevent Details screen (first screen of the MALL/LL event flows)

The rest of the screens ask about the following event information: digital and non-digital materials, Chinese skills and content, relation to RW and RL (according to our model),

Proceedings of the XXIst International CALL Research Conference

141

and feelings. When learners finish filling up the questions, they click on the submit button. They can neither view nor edit the submitted event to avoid any diary-inflicted effect, thus keeping its primary function as a data collection instrument.

The Admin Monitor module

Researchers and practitioners can track learners' events and evaluate them through the Track Learners' Activity screen in the Admin Monitor module (Figure 3). This screen presents the number of active and completed participants (diarists whose study has not ended yet or has already ended), the number of their submitted events, and their overall time in the program. These figures are helpful to monitor the number of participants and how engaged they are. The "Learners' Submitted Events" report, exported from this screen, presents the complete event information, including automatically calculated information such as the total number of events and each event type. Researchers and practitioners should use it during tracking and in the data analysis phase. Thus, the diary adheres to the principles of facilitating data tracking and analysis and ongoing filling in of diaries.

← Study A <	Submitted Events repor
Active (13) Completed (31)	Active and completed
orit111 change PVCIkspn0LgeGum02BTXANq00y03 3 events submitted Study Sun. 20/03/22 - Sat. 23/03/30 Study time elaped: 25 days	participants
Orit Researcher 4MxEFPRNYCZSq0GBwAe6deRq1JM2 0 events submitted Study Sat. 26/03/22 - Tue. 26/03/30 Study time elaped: 19 days	
orit a F2ZYEfY3NI/UUCWOPsMHtjQ5N6I1 5 events submitted Study Thu. 31/03/22 - Thu. 14/04/22 Study time elaped: 14 days	
orit b J613WfsnMZdLJJd8nj9pB1ayT9x1 0 events submitted Study: Thu: 14/04/23	Number of submitted events
Study time elaped: 14 days	Time in program
orit0401	
3 events submitted Study Fri. 01/04/22 - Fri. 15/04/22 Study time elaped: 13 days	
orit0405 RNpRXCTiaPqW5W/CuzGoNZFsql1 0 events submitted Study Tuie. 05/04/22 - Tuie. 19/04/22 Study Tuie. elaned: 9 days	



Discussion and conclusions

We developed the language diary app based on the DRM guideline (one-day-backward recollections) (Kahneman et al., 2004), eight design principles, and the RW/RL contextualized learning model developed in our past studies (Cohen & Ezra, 2018; Ezra & Cohen, 2018). These foundations make the diary usable in future studies and learning programs regardless of the target language and with little need for coding. As best as we know, this app is a forerunner pioneering dedicated diary application for language learning, specifically designed for the purposes of collecting data about outside-classroom RW-RL CLL & CMALL, that extends beyond previous language-learning electronic diaries (Benson et al., 2018; García-Amaya, 2017; Ranta & Meckelborg, 2013). The Learner Diary module's reporting events functionality is based on the RW/RL

contextualized learning model adjusted to the diary. Diarists are guided unequivocally, based on this model, using clear instructions and questions about aspects of language, context, materials, and feelings. Thus, this diary is more structurally consistent and comprehensive than traditional language-learning diaries, both open-ended ones (e.g., Benson et al., 2018; Litzler, 2014b; Litzler & Bakieva, 2017a; Warden et al., 1995) and more structured diaries (X.-B. Chen, 2013; García-Amaya, 2017; Hyland, 2004; Krishnan & Hoon, 2002; Ranta & Meckelborg, 2013). The diary questions are easily configurable and require no coding, allowing for a swift transition from one program to the next. The Admin Monitor module enables the user to track learners' progress and analyze the data efficiently and comprehensively. We plan in the next stage to perform comprehensive QA tests followed by several data-collection rounds. These rounds will be held in Taiwan among students of Chinese as a second language.

Language-learning diary studies and learning programs can draw on the eight design principles suggested in this paper as a reference checklist. Studies and learning programs interested in learning about outside-classroom RW-RL CLL & CMALL can draw assistance from data reported in the diary app. Furthermore, as developing materials in advance for contextualized learning requires knowledge of the entire domain's participants, situations, and language (Reinders & White, 2010), we suggest a data collection app such as Langy is also purposeful for material development. It informs about where, when, and what is learned outside class in RW-RL CLL & CMALL, supporting material adaptation to these outside-class learning contexts. In the future, technology-assisted reconstruction methods of moment assessment (Karapanos, 2020), using wireless communication, and sensing technologies (G. J. Hwang & Wu, 2014) can be used as well to provide a complementary picture of the phenomena. In MALL, tracking, including app records and server logs (Stockwell, 2022), can be integrated as well. Finally, we believe the diary-based instrument can be used in the future as a language-learning tool, although this will require additional development.

Acknowledgements

This research is supported by: 1. the Minducate – Science of Learning Research and Innovation Center. Sagol School of Neuroscience, Tel Aviv University, Israel. 2. Israel Polak Program for Excellence, Tel Aviv University, Israel.

References

- Benson, P., Chappell, P., & Yates, L. (2018). A day in the life: Mapping international students' language learning environments in multilingual Sydney. *Australian Journal of Applied Linguistics*, 1(1), 20–32. https://doi.org/10.29140/ajal.v1n1.21
- Bradburn, N. M., Rips, L. J., & Shevell, S. K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236(4798), 157–161. https://doi.org/10.1126/science.3563494
- Bradley, L., Lindström, N. B., & Hashemi, S. S. (2017). Integration and language learning of newly arrived migrants using mobile technology. *Journal of Interactive Media in Education*, 2017(1). http://doi.org/10.5334/jime.434
- Brown, M., Hughes, H., Keppell, M., Hard, N., & Smith, L. (2015). Stories from students in their first semester of distance learning. *The International Review of Research in Open and Distributed Learning*, *16*(4). https://doi.org/10.19173/irrodl.v16i4.1647
- Chanjavanakul, N. (2017). "Learning without intending to": A qualitative study of high school students' out-of-class English learning practices. [Unpublished doctoral dissertation]. University of California.
- Chen, C. M., & Li, Y. L. (2010). Personalised context-aware ubiquitous learning system for supporting effective English vocabulary learning. *Interactive Learning Environments*, *18*(4), 341–364. http://dx.doi.org/10.1080/10494820802602329

Proceedings of the XXIst International CALL Research Conference

143

Chen, X.-B. (2013). Tablets for informal language learning: Student usage and attitudes. Language Learning & Technology, 17(1), 20–36. http://dx.doi.org/10125/24503

Cohen, A., & Ezra, O. (2018). Development of a contextualised MALL research framework based on L2 Chinese empirical study. *Computer Assisted Language Learning*, *31*(7), 764–789. https://doi.org/10.1080/09588221.2018.1449756

Comas-Quinn, A., Mardomingo, R., & Valentine, C. (2009). Mobile blogs in language learning: Making the most of informal and situated learning opportunities. *ReCALL*, 21(01), 96–112. https://doi.org/10.1017/S0958344009000032

Curtis, A., & Bailey, K. (2009). Diary studies. OnCUE Journal, 3(1), 67–85.

Edge, D., Searle, E., Chiu, K., Zhao, J., & Landay, J. A. (2011). MicroMandarin: Mobile language learning in context. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3169–3178.

Ezra, O., & Cohen, A. (2018). Contextualised MALL: L2 Chinese students in target and non-target country. *Computers & Education*, 125, 158–174. https://doi.org/10.1016/j.compedu.2018.06.011

García-Amaya, L. (2017). Detailing L1 and L2 use in study-abroad research: Data from the daily linguistic questionnaire. System, 71(2017), 60–72. http://dx.doi.org/10.1016/J.SYSTEM.2017.09.023

Hwang, G. J., & Wu, P. H. (2014). Applications, impacts and trends of mobile technology-enhanced learning: A review of 2008-2012 publications in selected SSCI journals. *International Journal of Mobile Learning and Organisation*, 8(2), 83–95. https://doi.org/10.1504/IJMLO.2014.062346

Hwang, W. Y., Chen, H. S. L., Shadiev, R., Huang, R. Y. M., & Chen, C. Y. (2014). Improving English as a foreign language writing in elementary schools using mobile devices in familiar situational contexts. *Computer Assisted Language Learning*, 27(5), 359–378. https://doi.org/10.1080/09588221.2012.733711

Hyland, F. (2004). Learning autonomously: Contextualising out-of-class English language learning. *Language Awareness*, *13*(3), 180–202. https://doi.org/10.1080/09658410408667094

Ibrahim, A., Huynh, B., Downey, J., Höllerer, T., Chun, D., & O'Donovan, J. (2017). ARbis Pictus: A study of language learning with augmented reality. *ArXiv Preprint ArXiv:1711.11243*.

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2004). A survey method for characterizing daily life experience: The day reconstruction method. *Science*, *306*(5702), 1776–1780. https://doi.org/10.1126/science.1103572

Karapanos, E. (2020). Technology-assisted reconstruction: A new alternative to the experience sampling method. *Behaviour & Information Technology*, 39(7), 722–740. https://doi.org/10.1080/0144929X.2019.1608303

Krishnan, L. A., & Hoon, L. H. (2002). Diaries: Listening to 'voices' from the multicultural classroom. *ELT Journal*, 56(3), 227–239. https://doi.org/10.1093/ELT/56.3.227

Kukulska-Hulme, A. (2012). Mobile-assisted language learning. In C. Chapelle (Ed.), *The* encyclopedia of applied linguistics (pp. 3701–3709). Wiley Online Library. https://doi.org/10.1002/9781405198431.wbeal0768.pub2

Lamb, M., & Arisandy, F. E. (2020). The impact of online use of English on motivation to learn. *Computer Assisted Language Learning*, 33(1–2), 85–108. https://doi.org/10.1080/09588221.2018.1545670

Larson, R., & Csikszentmihalyi, M. (2014). The experience sampling method. In M. Csikszentmihalyi (Ed.), *Flow and the foundations of positive psychology: The collected works of Mihaly Csikszentmihalyi* (pp. 21–34). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9088-8_2

Lee, S.-M. (2019). A systematic review of context-aware technology use in foreign language learning. *Computer Assisted Language Learning*. https://doi.org/10.1080/09588221.2019.1688836

Lee, S.-M., & Park, M. (2019). Reconceptualization of the context in language learning with a location-based AR app. *Computer Assisted Language Learning*, 33(8), 936–959. https://doi.org/10.1080/09588221.2019.1602545 Litzler, M. F. (2014a). Independent study logs: Guiding and encouraging students in the process of language learning. *Journal of Language Teaching and Research*, 5(5), 994–998. https://doi.org/10.4304/jltr.5.5.994-998

Litzler, M. F. (2014b). Student satisfaction with using independent study logs as part of English classes. *Theory and Practice in Language Studies*, 4(8), 1527–1535. https://doi.org/10.4304/TPLS.4.8.1527-1535

Litzler, M. F., & Bakieva, M. (2017a). Learning logs in foreign language study: Student views on their usefulness for learner autonomy. *Didáctica*, 29, 65–80.

Litzler, M. F., & Bakieva, M. (2017b). Students' out-of-class activities for practicing English as a foreign language: A log study. *International Journal of Scientific Study*, 5(8), 10–19. 10.17354/ijssNov/2017/2

Pegrum, M. (2014). *Mobile learning: Languages, literacies and cultures*. Palgrave Macmillan. https://doi.org/10.1057/9781137309815

Ranta, L., & Meckelborg, A. (2013). How much exposure to English do international graduate students really get? Measuring language use in a naturalistic setting. *The Canadian Modern Language Review*, 69(1), 1–33. http://dx.doi.org/10.3138/cmlr.987

Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. Language Teaching, 50(4), 561–578. https://doi.org/10.1017/S0261444817000192

Reinders, H., & White, C. (2010). The theory and practice of technology in materials development and task design. In N. Harwood (Ed.), *English language teaching materials: Theory and practice* (pp. 58–80). Cambridge University Press.

Rivers, D. J. (2009). Utilizing the quick response (QR) code within a Japanese EFL environment. *JALT CALL Journal*, *5*(2), 15–26.

Sandberg, J., Maris, M., & de Geus, K. (2011). Mobile English learning: An evidence-based study with fifth graders. *Computers & Education*, 57(1), 1334–1347. https://doi.org/10.1016/j.compedu.2011.01.015

Santos, M. E. C., Lübke, A. in W., Taketomi, T., Yamamoto, G., Rodrigo, M. M. T., Sandor, C., & Kato, H. (2016). Augmented reality as multimedia: The case for situated vocabulary learning. *Research and Practice in Technology Enhanced Learning*, 11. https://doi.org/10.1186/s41039-016-0028-2

Sharples, M., Taylor, J., & Vavoula, G. (2007). A theory of learning for the mobile age. In R. Andrews & C. A. Haythornthwaite (Eds.), *The Sage handbook of e-learning research* (pp. 221–247). Sage.

Steel, C. H., & Levy, M. (2013). Language students and their technologies: Charting the evolution 2006–2011. *ReCALL*, 25(3), 306–320. http://doi.org/10.1017/S0958344013000128

Stockwell, G. (2022). Mobile assisted language learning: Concepts, contexts and challenges. Cambridge University Press. https://doi.org/10.1017/9781108652087

Vavoula, G. N. (2005). A study of mobile learning practices. University of Leicester. https://leicester.figshare.com/ndownloader/files/18268073

Viberg, O., & Grönlund, Å. (2012). Mobile assisted language learning: A literature review. 11th World Conference on Mobile and Contextual Learning, MLearn 2012.

Viberg, O., Khalil, M., & Bergman, G. (2021). TimeTracker app: Facilitating migrants' engagement in their second language learning. In M. E. Auer & T. Tsiatsos (Eds.), *IMCL 2019. Advances in intelligent systems and computing* (Vol. 1192, pp. 983–994). Springer, Cham.

Warden, M., Lapkin, S., Swain, M., & Hart, D. (1995). Adolescent language learners on a three-month exchange: Insights from their diaries. *Foreign Language Annals*, 28(4), 537–550. https://doi.org/10.1111/J.1944-9720.1995.TB00827.X

Zhang, S. (2018). Augmented reality in foreign language education: A review of empirical studies. Journal of Technology and Chinese Language Teaching, 9(2), 116–133.

Proceedings of the XXIst International CALL Research Conference

145

Paul John, Walcir Cardoso, Carol Johnson

Université du Québec à Trois-Rivières, Trois-Rivières, Canada Concordia University, Montreal, Canada Concordia University, Montreal, Canada

paul.john@uqtr.ca, walcir.cardoso@concordia.ca, carol.johnson@concordia.ca

On the adequacy of L2 pronunciation feedback from automatic speech recognition: A focus on Google Translate

Bio data



Paul John is Associate Professor of Modern Languages at the Université du Québec à Trois-Rivières. His research focuses on L2 phonological acquisition, with particular interest in variation and in using neuroimaging to investigate L2 perception. Another research interest concerns the use of automatic speech recognition, text-to-speech and grammar checkers for language learning.



Walcir Cardoso is Professor of Applied Linguistics at Concordia University. He conducts research on the L2 acquisition of phonology, morphosyntax and vocabulary, and the effects of computer technology (e.g., clickers, text-to-speech synthesizers, automatic speech recognition, intelligent personal assistants) on L2 learning.



Carol Johnson is a PhD student in Education (specialization in Applied Linguistics) at Concordia University. Her research interests include the pedagogical use of speech technologies in L2 learning, particularly the use of automatic speech recognition to improve pronunciation and writing. She currently teaches university-level ESL courses.

Abstract

This study investigates automatic speech recognition (ASR) in Google Translate as a source for L2 pronunciation feedback. To be effective, ASR should transcribe learner errors accurately and perform equally well on male and female voices, avoiding gender bias. We assess Google Translate on three Quebec francophone (QF) segmental errors in English: th-substitution (*think* \rightarrow [t]*ink*); h-deletion (*happy* \rightarrow _*appy*); and h-epenthesis (*ice* \rightarrow [h]*ice*). Eight QFs (4F/4M) recorded 120 sentences with and without an error on the final item (e.g., *I don't know who to *tank/thank*). Errors were equally divided between real word output (**tank*) and nonword output (e.g., *My sister is afraid of *tunder*). We anticipate real word errors, corresponding to entries in the Google Translate lexicon, will be accurately transcribed, whereas nonwords, by definition absent from the lexicon, should be erroneously matched to similar-sounding real words (i.e., the intended output "thunder"), constituting misleading feedback.

Forthcoming data analyses will determine the relative contribution of error type, real/nonword output, and gender to final-word transcription and feedback accuracy. Preliminary findings suggest a hierarchy of accuracy (h-deletion, h-epenthesis >

Proceedings of the XXIst International CALL Research Conference

th-substitution) specific to real-word output. Indeed, ASR shows a clear inability to flag nonword errors. A gender bias effect is not apparent; in fact, ASR generally transcribed the sentences recorded by females more accurately. Mistranscriptions unrelated to final items have yet to be examined. Our presentation will address the implications of our findings for L2 teachers/learners and for developers seeking to design ASR specifically for L2 uses.

Conference paper

Introduction

Second language (L2) learners generally require help with pronunciation since L2 phonology is hard to acquire through mere exposure (Flege, Munro & Mackay, 1995). One means of promoting pronunciation accuracy in a classroom context is through corrective feedback (Lyster et al., 2013). The potential of automatic speech recognition (ASR) to supply feedback that learners can access autonomously and ubiquitously is thus appealing. ASR is a function widely available in tools such as Google Translate (GT), the system focused on here. Nonetheless, different aspects of the adequacy of ASR feedback remain to be determined.

The current study investigates ASR feedback with respect to three segmental pronunciation errors typical of Quebec francophone (QF) learners of English: th-substitution (*thank* \rightarrow *tank*), h-deletion (*heat* \rightarrow *_eat*) and h-epenthesis (*ice* \rightarrow *hice*) (Brannen, 2011; John & Cardoso, 2009; Mah et al., 2016; White et al., 2015). These errors are usually variable. QF learners alternate between pronouncing items such as *thank*, *heat* and *ice* inaccurately and accurately: [tæŋk]~[θ æŋk]; [it]~[hit]; and [hajs]~[ajs]. For pronunciation feedback purposes, ASR systems should distinguish between correct and incorrect pronunciation of L2 segments, reinforcing learners' accurate production while also flagging inaccuracies.

First, when QFs pronounce *thank*, *heat* and *ice* correctly, it is crucial that ASR confirms the targetlike output by providing an accurate transcription. Otherwise, by transcribing something else the system is sending the misleading message that learners' pronunciation is off (a 'false alarm'). Second, when QFs mispronounce *thank*, *heat* and *ice*, it is important that ASR indicates that learners have missed the mark. Importantly, we anticipated that such corrective feedback would be more accurate with mispronunciations leading to real-word output (e.g., *thank* \rightarrow *tank*) than nonword output (*ice* \rightarrow *hice*).

The reason is that ASR bases its transcriptions on items stored in its lexicon. ASR attempts to match the phonetic output to existing lexical entries. Consequently, the system should transcribe learner errors more accurately when the output corresponds to a real word, directly signaling that learners have mispronounced the target item. In the case of nonword output (*thief* \rightarrow *tief*, *head* \rightarrow *_ead*, *ice* \rightarrow *hice*), ASR will likely search for the closest lexical match for the phonetic output, potentially arriving at *thief*, *head* and *ice*. This is particularly the case when items are embedded in sentences, such that ASR can make use of top-down prediction based on collocations, semantic associations and syntactic analysis. With correctly pronounced items, top-down processing helps the system make an accurate lexical match, thus reducing the likelihood of false alarms. With incorrectly pronounced items, however, top-down processing risks overriding phonetic cues, supplying the appropriate (i.e., target) item for the context despite the mispronunciation. To our knowledge, these issues have yet to be investigated empirically.

Our study examines ASR transcriptions for accurately and inaccurately pronounced target items that begin with $/\theta/$, /h/ or a vowel. One objective was to establish the extent to which, in a sentence context, ASR generates false alarms when faced with accurately pronounced items. Furthermore, we used inaccurately pronounced items constituting

real-word or nonword output. We anticipated the following hierarchy for transcription accuracy:

no error condition > error condition (real-word) > error condition (nonword)

Given mispronounced items, another objective was to explore whether ASR flags the three error types with comparable accuracy. In addition, the sentences were recorded by male and female QF learners of English in order to investigate the question of gender bias. Given that ASR systems are often trained on corpora with a preponderance of male voices, they may perform less well on female output (Tatman, 2017). For L2 teachers, it would be a serious concern if they were to recommend a tool that provides less accurate feedback to their female learners.

Methodology

Materials

148

A total of 120 sentences with a final item starting with either $/\theta/$, /h/ or a vowel were recorded by 4 male and 4 female QFs with and without a pronunciation error on the final item. For example:

I don't know who to **thank/*tank.** I still need to brush my **hair/*air**. We slipped and fell on the **ice/*hice**.

In the error condition, the 120 sentences were evenly divided between 60 real-word and 60 nonword errors.

Data collection and analysis

480 recordings (4 versions of each sentence: in the error and no error conditions spoken by a male and female speaker) were played into GT's ASR function. We noted whether each transcription corresponded to the *error* or *no error* condition and whether the speaker was *male* or *female*. We further noted whether the final target item began with $/\theta/$, /h/ or a vowel and whether, given an error, the output contained a *real-word* or *nonword*. The following independent variables were thus included: pronunciation accuracy (error-no error), gender (M-F), target sound ($/\theta/-/h/-V$), and output form in the error condition (real-nonword).

The ASR transcriptions of final items were also examined for two dependent variables: *transcription accuracy* and *accuracy of pronunciation feedback*. Transcription and feedback accuracy are partly, but not entirely, coextensive. Where a transcription captures exactly the phonetic output, clearly it provides accurate pronunciation feedback. Conversely, if target *thank* is correctly pronounced [Øæŋk] but transcribed 'tank' OR if target *thank* is mispronounced [tæŋk] but transcribed 'thank', this constitutes incontrovertibly inaccurate feedback on pronunciation.

A transcription can, however, diverge from these clearcut cases (e.g., $[\theta @n]k]$ may be transcribed 'talk' or 'thong' or 'sank'), and the resulting feedback can vary in how misleading it is. We classified such inaccurate transcriptions into *problematic*, *partaccurate* and *neutral feedback*.

In the no error condition, given a target with initial / θ /, /h/ or a vowel, the categories refer to:

Problematic feedback: The transcribed item starts with /t/, a vowel or /h/, sending the misleading message that the learner has engaged in th-substitution, h-deletion or h-epenthesis.

Part-accurate feedback: Although different from the target item, the transcribed item starts with $\theta/$, h/ or a vowel, appropriately signaling the correct pronunciation.

Neutral feedback: The transcribed item starts with another sound entirely, indicating neither a mispronunciation nor a correct pronunciation.

In the error condition, *problematic* and *part-accurate feedback* are the reverse of above; *neutral feedback* is the same:

Problematic feedback: The transcribed item starts with $\theta/$, h/ or a vowel, thus masking the mispronunciation of the initial segment.

Part-accurate feedback: Although different from the target item, the transcribed item starts with /t/, a vowel or/h/, appropriately signaling the mispronunciation.

Neutral feedback: The transcribed item starts with another sound entirely, neither masking nor correctly signaling the mispronunciation.

Results

No error condition

Accuracy rates for transcriptions of correctly pronounced final items provide an indication of the extent to which the ASR system confirms correct pronunciation, thus providing accurate feedback. Table 1 shows the transcription accuracy rates for final items that, were they mispronounced, would result in a) real-word or b) nonword output, for M and F speakers both separately and combined. The overall mean for both sets of sentences (a + b) is also indicated.

-	•	-	-
Target items	м	F	M + F
th-initial	.70	.80	.75
h-initial	.85	.95	.90
V-initial	.90	.95	.93
Mean:	.82	.90	.86

Table 1. No error condition: final word transcriptions

a. Accuracy rates (real-word output sentences in error condition)

b. Accuracy rates (nonword output sentences in error condition)

Target items	М	F	M + F
th-initial	.75	.90	.83
h-initial	.95	1.0	.98
V-initial	.85	1.0	.93
Mean:	.85	.95	.90
Overall mean:	.83	.93	.88

Across the board, transcription accuracy rates for F voices are higher than for M voices (Table 1), the opposite to the predicted pattern of gender bias. In terms of the target

Proceedings of the XXIst International CALL Research Conference

items, h-initial and V-initial items have consistently higher accuracy rates than th-initial items. Accuracy rates were incidentally higher in the set of sentences that would lead to nonword output in the error condition (except vowel-initial items spoken by M voices). Conceivably, the final items in this second set of sentences are marginally more predictable.

	PRONUNCIATION FEEDBAC	CK		
Target items	Inaccurate (<i>thank</i> \rightarrow <i>tank</i> , etc.)	Problematic (theft \rightarrow test)	Part-accurate (<i>thrifty</i> \rightarrow <i>thirsty</i>)	Neutral $(thud \rightarrow fun)$
th-initial	.00	.03	.05	.14
h-initial	.00	.00	.06	.00
V-initial	.00	.00	.06	.01

 Table 2. No error condition: rates of inaccurate transcriptions/feedback types

Likewise, very few of the mistranscriptions indirectly send the message that the learner has produced a typical pronunciation error. In fact, the only examples of such problematic feedback involve 2 sentences with target items starting with $/\theta/$ (e.g., *I* wonder who committed the **theft** \rightarrow **test**). The mistranscriptions for 4 sentences with target items starting with $/\theta/$ in fact substitute items that include $/\theta/$, thus constituting part-accurate feedback (e.g., *Danielle has always been very* **thrifty** \rightarrow **thirsty**). While strictly speaking inaccurate, these mistranscriptions nonetheless indirectly suggest learners have correctly realized the difficult target sound.

All of the mistranscriptions of sentences with target items starting with /h/ and all but one with target items starting with a vowel were of this same part-accurate feedback type (Table 2). No examples of problematic or neutral feedback were observed for h-initial forms; no examples of problematic feedback were observed for vowel-initial forms. The item transcribed in each case started with /h/ or a vowel (e.g., *Their wedding was in a lovely hall* \rightarrow *whole*; *You need to use your elbows* \rightarrow *albums*). With such part-accurate feedback, learners will not be inclined to conclude they have deleted /h/ or needlessly inserted it.

The remaining 11 mistranscriptions of items starting with θ constitute neutral feedback (i.e., the transcription contained another sound than θ or t: *The book hit the floor with a loud thud* \rightarrow *fun*). One mistranscription of a target item beginning with a vowel also constitutes neutral feedback: *Chicken pox makes you itch* \rightarrow *pitch*.

Error condition

150

Accuracy rates for transcriptions of mispronounced final items provide an indication of the extent to which ASR flags pronunciation errors, thus providing accurate feedback. Table 3 shows the rates of accurate transcription for mispronounced items corresponding to a) real-word or b) nonword output.

Table 3. Error condition: final word transcriptions

Target items	М	F	M + F
th-initial	.20	.30	.25
h-initial	.35	.60	.48
V-initial	.30	.85	.58
Mean	27	58	43

a. Accuracy rates (real-word output)

b. Accuracy rates (nonword output)

Target items	М	F	M + F
th-initial	.00	.00	.00
h-initial	.05	.15	.10
V-initial	.05	.20	.13
Mean:	.03	.12	.08
Overall mean:	.15	.35	.26

Transcription accuracy rates for F voices are, as in the error condition, almost always higher than for M voices (Table 3). This F advantage is the opposite to the predicted pattern of gender bias. H-initial and vowel-initial items once again show consistently higher accuracy rates than th-initial items. Accuracy rates are also decidedly lower in the set of sentences with nonword output. Indeed, as expected, none of the mispronounced th-initial targets resulting in nonwords were correctly transcribed. Surprisingly, some h-initial targets (1M and 3F) and some vowel-initial targets (1M and 4F) resulting in nonwords were correctly transcribed. In some cases, GT found a proper noun that corresponded to the supposed nonword (e.g., for oil \rightarrow hoil, the transcription was 'Hoyle'). In one sentence, empty \rightarrow hempty was transcribed 'hemp tea'.

The overall mean of .26 for M and F voices suggests learners can expect ASR to mistranscribe 7.4 of 10 content words that they mispronounce in a sentence context. Nonetheless, transcription accuracy varies considerably: while accuracy rates for nonword output are exceedingly low (as low as .00), rates for real-word output can be promisingly high (as high as .85). Correspondingly, the degree of truly accurate pronunciation feedback also varies.

Incorrect transcriptions that provide clearly inaccurate pronunciation feedback are those that supply the mispronounced target item (e.g., $thank \rightarrow tank$ is transcribed 'thank'). As reported above, no examples of the reverse phenomenon ($thank \rightarrow tank$ being transcribed 'tank') appear in the no error condition. In the error condition, however, such clearly inaccurate feedback is widespread. Table 4 shows the rates of target item transcription for mispronounced final items corresponding to a) real-word or b) nonword output.

Target items	м	F	M + F	
th-initial	.50	.45	.48	
h-initial	.40	.30	.35	
V-initial	.40	.15	.28	
Mean:	.43	.30	.37	

Table 4. Error condition: target item transcriptions (inaccurate feedback)

b. Target item rates (nonword output sentences)

a Target item rates (real-word output sentences)

Target items	М	F	M + F
th-initial	.65	.70	.68
h-initial	.65	.70	.68
V-initial	.65	.55	.60
Mean:	.65	.65	.65
Overall mean:	.54	.47	.51

While target items are erroneously transcribed just over a third of the time with real-word output, this rises to almost two thirds with nonword output. This is precisely the pattern of transcription inaccuracy – and concomitant feedback inaccuracy – according to the output type that we anticipated.

Nonetheless, not all of the mistranscriptions supplied the target item, constituting inaccurate feedback. Instead, some mistranscriptions supplied items other than the target, constituting problematic, part-accurate or neutral feedback (Table 5).

	PRONUNCIATION FEEDBACK				
Target items	Inaccurate $(tank \rightarrow thank, etc.)$	Problematic $(tumb \rightarrow thong)$	Part-accurate (<i>teft</i> \rightarrow <i>test</i>)	Neutral $(tud \rightarrow pop)$	
th-initial	.58	.03	.20	.08	
h-initial	.51	.00	.14	.05	
V-initial	.44	.00	.20	.01	

Table 5. Error condition: rates of inaccurate transcriptions/feedback types

Problematic feedback comes from ASR transcribing other items that start with θ /, /h/ or a vowel, given target items with these sounds. Only two such inaccurate transcriptions occurred. Both involve th-initial targets (e.g., *Unfortunately, she twisted her tumb* \rightarrow *thong*). Inaccurate transcriptions that provide this problematic form of pronunciation feedback are thus extremely rare.

Most of the mistranscriptions resulting in items other than the targets generated part-accurate feedback that indirectly captures the pronunciation error. Indeed, 16 of the

Proceedings of the XXIst International CALL Research Conference

23 mistranscriptions of this sort given a th-initial target provide t-initial items, reflecting the substitution error (e.g., *I wonder who committed the* **teft** \rightarrow **test**). This is likewise the case for 11 of 15 mistranscriptions given h-initial targets and for 16 of 17 mistranscriptions given vowel-initial targets: the items are vowel-initial and h-initial respectively, reflecting the h-deletion and h-epenthesis errors.

The few remaining mistranscriptions involving items other than the target item constitute neutral feedback. That is, the transcribed item contains another sound entirely (e.g., *The book hit the floor with a loud* **tud** \rightarrow **pop**). This is the case for 6 of 23 mistranscriptions of th-initial targets, for 4 of 15 mistranscriptions of h-initial targets, and for 1 of 17 mistranscriptions of vowel-initial targets.

Mistranscriptions of inaccurately realized items are thus often of the false negative type: they erroneously indicate that the speaker has correctly realized the target item. This is particularly the case when the phonetic output constitutes a nonword. Mistranscriptions that supply items other than the target item, however, in most cases supply part-accurate feedback: the transcription provides an item that captures the mispronunciation of the difficult sound, thus indirectly signaling the pronunciation error.

Discussion/Conclusion

To recap, given accurately pronounced items (no error condition) in a sentence context, ASR in GT is highly unlikely to generate false alarms that would send L2 learners the misleading message that they have made a pronunciation error. Indeed, across 240 correctly realized sentences, only two mistranscribed target items created this false impression. This is highly reassuring from a corrective feedback perspective. Nonetheless, for feedback purposes, it is even more important how ASR transcribes inaccurately pronounced items. Unfortunately, ASR struggled to transcribe pronunciation errors accurately, particularly given nonword (.08) vs real-word (.43) output. The tendency was to transcribe the contextually appropriate target item (overall: .51) rather than the phonetically accurate item, especially given nonword output (.65). Nonetheless, the higher transcription accuracy for some items (notably .85 for F vowel-initial forms) suggests that, even given a sentence context, ASR can at times provide effective feedback. In addition, when ASR transcribed an item other than the output or target item, in most cases the initial sound was captured in the transcription, constituting part-accurate feedback. Teachers and learners who want to target th-substitution errors should be forewarned that ASR experienced greater difficulty correctly transcribing th-initial than h-initial or vowel-initial items, whether in the no error or error condition. The concern that ASR might show a gender bias, however, is not supported: in fact, ASR generally showed higher accuracy when transcribing the female recordings.

In sum, our impression is that ASR probably provides better pronunciation feedback on items spoken in isolation that avoid the influence of contextual cues and especially on minimal pairs that avoid the nonword transcription problem. Future research will investigate this issue further.

References

- Brannen, K. (2011). The perception and production of interdental fricatives in second language acquisition [Doctoral dissertation, McGill University].
- Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. Speech Communication, 16(1), 1–26.
- Lyster, R., Saito, K., & Sato, M. (2013). Oral corrective feedback in second language classrooms. *Language Teaching*, 46(1), 1-40.

Proceedings of the XXIst International CALL Research Conference

- John, P., & Cardoso, W. (2009). Francophone ESL learners' difficulties with English /h/. In B. Baptista, A. Rauber & M. Watkins (Eds.), *Recent research in second language phonetics/phonology: Perception and production* (pp. 118-140). Cambridge.
- Mah, J., Goad, H., & Steinhauer, K. (2016). Using event-related brain potentials to assess perceptibility: The case of French speakers and English [h]. Frontiers in Psychology, 7, 1-14.
- Tatman, R. (2017). Gender and dialect bias in YouTube's automatic captions. In D. Hovy, S. Spruit, M. Mitchell, E. M. Bender, M. Strube & H. Wallach (Eds.), *Proceedings* of the First Workshop on Ethics in Natural Language Processing (pp. 53-59). Association for Computational Linguistics.
- White, E. J., Titone, D., Genesee, F., & Steinhauer, K. (2015). Phonological processing in late second language learners: The effects of proficiency and task. *Bilingualism:* Language and Cognition, 1-22.

Phuong Tran

Waseda University, Tokyo, Japan

phuongtran@fuji.waseda.jp

Effect of English language proficiency on learner engagement in communities of inquiry

Bio data



Dr. Phuong Tran, Lecturer at Waseda University, has taught English and self-directed learning courses in Austria, Japan and Vietnam for over ten years. Her research investigates the development of learner autonomy through social interaction and networking in language learning settings. She has written a number of academic articles and book chapters in applied linguistics.

Abstract

In educational settings, social networking may be thought as being closely related to concepts such as teaching presence, social presence, and cognitive presence as a part of communities of inquiry (CoI) (see Garrison et al., 2000; Garrison et al., 2010; Swan & Shea, 2004), and how these relate to learner engagement in both the community itself and in tasks and activities that are the object of discussion in the community. Furthermore, learners' learning outcomes can be implied through engagement which is considered as a multidimensional concept that comprises cognitive, behavioural, social, and emotional dimensions (see Philp & Duchesne, 2016). As in current research from Stockwell (2019), how to maintain active participation in language learning tasks depends on what task engagement is. The types of engagement among second and foreign language learners in the classroom or outside of class may be different according to the level of proficiency, which forms the main focus of the study.

Conference paper

Introduction

The purpose of the study is to explore the similarities and differences of learner engagement between students with low and high proficiency levels of English through social networking with the assistance of teacher presence, and to determine how social networking and teacher presence can encourage task engagement according to the level of learner proficiency outside of class. The guiding key scientific question for the project is (a) whether the interaction by the teacher with different types of learners through social networking can promote different types of engagement in tasks outside the classroom, and (b) whether there are any differences and similarities in types of engagement among learners of English. With the advances of digital and powerful online tools with which learners can construct their understanding of content through guidance and support from peers and instructors, the venue potentially serves the educational community well (Meskill, 2013), as well as with the continuous increase of using

multimodalities for language teaching, which increase educational opportunities. However, there has been little research on how these uses support student engagement in instructional activities (Kozlova & Zundel, 2013), especially support on different types of learner proficiency with different types of engagement.

Moreover, throughout my recent studies, there are three main terms worth noticing, which lead to my currently proposed project: (1) in terms of the role of teacher presence, based on the theory of Community of Inquiry (CoI) (Garrison et al., 2000), teacher presence had all three elements including social presence, cognitive presence and teaching presence. From the findings, the first element is the teacher's social presence having the same categories as the theory suggested consisting of open communication, group cohesion and affective expression. The teacher's role is to start the conversation, balance the interaction and give encouragement to the whole class. Similarly, the second element is the teacher's cognitive presence with the same characteristics mentioned in CoI including triggering event, exploration, integration and resolution. The teacher encouraged students by raising questions, exploring their strengths and weaknesses and trying to figure out a solution for students' learning. The last element is the teacher's teaching presence including design and organization, facilitating discourse and direct instruction, however, we added one more characteristic to the original theory: reminder. The teacher showed her experiences and skills in guiding students on how to do the tasks in general, giving feedback and corrections as well as advice. The teacher also showed her humble side as a reminder, simply in promoting task engagement among learners. (2) With regard to the role of online social interaction, as mentioned in Tran (2018), there are three main roles of social interaction. The first role is as 'a social community discussion tool' (p.) where learners can have friendly conversation and discussion with the instructor and group members. It is called a 'social tool' because of its features which allow unconstrained dialogues among interlocutors. Considering the low proficiency level of English of participants in the study, this free and easy interaction may have served as a motivating medium for them to use English in a non-threatening environment. Also, they can feel a sense of belonging to a group that they can share their learning with. The second role is as 'teacher-to-student reporting tool.' The teacher can use this social interaction as a tool to share their expertise, for example, sending extra materials, giving instructions and guidance, sending reminders and announcements, providing feedback and advice, and so on.

Although this may raise some arguments as creating this online social interaction tool outside of class means the teacher has to dedicate their time for students outside class time as well. This could be a time consuming burden, however, the good side of this setting is that it would serve as a fundamental stepsfor students to build their habits of learning outside of classroom which may ultimately lead to autonomous learning behaviour. The third role is as a 'student-to-teacher reporting tool.' The students use this online platform to communicate with the teacher, report what they have done outside of class or seek advice in a friendly way. (3) As regards impacts on task engagement, Hong (2008) claimed there is a close relationship between teachers' language use and students' engagement levels in the classroom. In my studies, the teacher's language use is present even outside of the classroom through the online social interaction tool, leading to nearly three times the level of student engagement in the designed tasks as well. Moreover, Van den Branden (2016) emphasized that the teacher plays a major role in motivating students through well-designed tasks that are both challenging and closely matched to their needs. In this sense, the teacher in my studies organized the listening tasks and vocabulary tasks for students in listening classes. The teacher also acknowledged the need to involve students through tasks that are strongly suitable and achievable with her support outside of the formal classroom setting.

As described above, the purpose of the research project is to identify the ways in which teacher presence through social networking can be used to enhance out-of-class learning by English learners of different proficiency levels, and to figure out whether there is a

relationship between the types of engagement and the types of learners with different proficiency levels of English. Research to date has been able to identify the existence of different types of teacher presence (Lee, 2014; Tran, 2018) as well as different types of engagement in the language classroom (Philp & Duchesne, 2016). However, there is very little research on understanding the relationship between teacher presence and task engagement, and the comparison of types of engagement among different levels of learners, which is believed to help teachers foster appropriate task engagement according to learner levels outside of class and lead to successful teaching and learning outcomes.

This is a significant advance beyond previous research which was only able to specify that social, teaching, and cognitive aspects can occur through social interaction in common learning communities, but this does not provide sufficient insight into how this knowledge can be applied to classroom practice. In other words, although there has been research determining that the role of the teacher is crucial in facilitating learning outside of class, there has been no published research thus far that explores different aspects of the teacher including social, teaching and cognitive presence through online social interaction. The project seeks to explore the potential techniques that teachers may be able to try through online social interaction with students, to determine: (1) How do low-level proficiency learners perceive social, teaching and cognitive presence of the teacher through the interaction? (2) How do high proficiency learners perceive the social, teaching and cognitive presence have on out-of-class task engagement among learners with different levels of proficiency and ultimate learning goals?

Data collection methods will be based on the methods that were previously adopted by Tran (2016, 2018), but will go beyond this in order to specifically compare the differences and similarities among low- and high-level proficiency learners with different types of engagement in online interaction through social networking. The study will be longitudinal in nature, consisting of a quasi-experimental method across two intact beginning English classes at a university in Japan with a total of around 55 participants, and two intermediate English classes at another private university in Japan with a total of around 55 participants, totaling 110 participants. The study will be carried out in two phases, the first of a more exploratory nature in 2021, and the second seeking to validate the results from the first phase in 2022. The selection of social networking tool for the classes is LINE for Japan, which is used by 100% of the learners in the classes the researcher has access to, and from previous study by the researcher is the preferred tool of learners.

Phase 1

For low proficiency learners of English

The teacher supported the learners with teaching, social and cognitive presence through the designated tool to assist the learners in engaging in English vocabulary and listening activities outside of class. Learner training on how to study with the vocabulary app and listening website were provided in class and using the tools throughout the research period. Data were collected through discourse analysis of the tool interaction logs to track what types of engagement the learners have in the interactions with class and individual preference. The logs enabled the teacher to figure out the trend of types of engagement in the tool interaction in terms of cognitive, behavioural, social, and emotional dimensions of engagement, and to see if this trend continued across the 15-week research period. In addition, weekly quizzes were conducted across the research period to compare engagement with outcomes from the quizzes to explore any potential correlation. On conclusion of the project, learners were administered a post-treatment reflections on their own engagement. Interviews were conducted with learners focusing on four different aspects of engagement, including cognitive, behavioural, social, and emotional to further explore their survey responses. The results were then collated to create a taxonomy of engagement types with a view to providing the teacher with specific learning strategies depending on the engagement types that low proficiency learners possess for promoting tasks outside of class.

For high proficiency learners of English

The teacher provided learners with support with teaching, social and cognitive presence through LINE to assist the learners in engaging in presentation contents and listening activities outside of class. Data were collected through discourse analysis of the tool interaction logs to track what types of engagement the learners have in the interactions with a common class group, small groups (4-5 members for each group) and individual interactions. Learner training on what to do with their presentation and how to study with the listening website were provided in class and through LINE across the research period. The logs enabled the teacher to figure out the trends of types of engagement in LINE interaction in terms of the cognitive, behavioural, social, and emotional dimensions of engagement, and to see if this trend continued across the 15-week research period. In contrast to the low proficiency group, in the high proficiency group, data on the content of small group presentations were analysed, and how they engaged with their presentation projects through LINE interactions outside of class were collected to explore any potential correlation with the interactions with peers and the teacher. On conclusion of the project, learners were administered a post-treatment questionnaire to see if they noticed the different types of engagement and their reflections on their own engagement. Interviews were conducted with learners on four different aspects of engagement including cognitive, behavioural, social, and emotional to further explore their survey responses. The results were then collated to create a taxonomy of engagement types with a view to providing the teacher with specific learning strategies depending on the engagement types that high proficiency learners possess for promoting tasks outside of class.

Phase 2

Data collection methods were mostly conducted in the same way as Phase 1, but the questionnaire and interviews conducted at the end of this phase explored whether the type of interactions had an impact on not only their types of engagement in the activities, but also on their attitudes towards the materials, towards learning English, and towards the communities of inquiry. In addition, data collected from Phase 1 were analysed using contrast analysis to compare the types of engagement for two different categories of learners to determine the differences and similarities in types of engagement and their correlation with task engagement. Low proficiency learners will not be assigned into small groups for their projects; however, high proficiency learners were required to work on their presentation projects, which is one of the main differences in the tasks assigned. The results were analysed to determine whether there were differences and similarities in how the learners interacted with one another, how they engaged with the task, and their view the learning environment. This resulted in a list of specific strategies that the research project identified as promoting task engagement depending on learners' level of proficiency and attitudes towards learning English outside of class through social networking.

In order to achieve the target results as well as identify perceptions of social networking, usage, attitudes and task engagement, surveys were distributed for all participants (N=110). Along with the pre- and post- surveys, the study also collected interview data with the post-treatment focus group, vocabulary engagement logs, quiz scores, and analysed all the LINE interactions in both individual and group chats.

After collecting data from the two phases, contrast analysis was used to analyse the

differences and similarities among Vietnamese learners of English and Japanese learners of English. The results were analysed to determine whether the culture that learners belong to affects types of engagement, and this resulted in a list of specific strategies for promoting task engagement with different types of learners.

References

- Garrison, D. R., Cleveland-Innes, M., & Fung, T. S. (2010). Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *Internet and Higher Education*, *13*, 31-35.
- Kozlova, I., & Zundel, E. (2013). Synchronous online language teaching: strategies to support learner development. In C. Meskill (Ed.), Online teaching and learning: Sociocultural perspectives (pp. 99–117). Bloomsbury Publishing.
- Lee, S. M. (2014). The relationships between higher order thinking skills, cognitive density, and social presence in online learning. *Internet and Higher Education*, 21, 41-52.
- Meskill, C. (Ed.). (2013). *Online teaching and learning: Sociocultural perspectives.* Bloomsbury Publishing.
- Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. Annual Review of Applied Linguistics, 36, 50-72.
- Stockwell, G. (2019). Insights from replication on the factors affecting task engagement in mobile-based learning activities. *Technology in Language Teaching & Learning*, 1(1), 33-35.
- Swan, K., & Shea, P. (2004). The development of virtual learner communities. In S. R. Hiltz & R. Goldman (Eds.), *Learning together online: Research on asynchronous learning networks* (pp. 239-260). Lawrence Erlbaum Associates.
- Tran, P. (2016). Training learners to use Quizlet vocabulary activities on mobile phones in Vietnam. *The JALT CALL Journal*, *12*(1), 43-56.
- Tran, P. (2018). Enhancing online language learning task engagement through social interaction. *Australian Journal Applied Linguistics*, 1(2), 88-101.

Sangmin-Michelle Lee

Kyung Hee University, Yongin, Korea

sangminlee@khu.ac.kr

Pedagogical benefits of technological affordances in a user-created metaverse space

Bio data



Dr. Sangmin-Michelle Lee is a Professor of Global Communication at Kyung Hee University in Korea. She earned her PhD from the Pennsylvania State University in language education. She has published papers and books on language learning in a technology-enhanced learning environment, machine translation, context-aware technologies, L2 writing, and digital creativity.

Abstract

Contextualization and interaction are critical to successful L2 learning. However, in English-as-a-foreign-language (EFL) situations where students do not have opportunities to use the target language, contextualized learning and interaction in English is seldom available and learning English is often limited to classroom practices. This leads to students' lack of interest and motivation to learn English. In recent years, as diverse virtual technologies advance, they offer a new learning environment for EFL students. In the so-called metaverse, spaces and situations can be created for learning English effectively. The current study explored potentials of using the metaverse space in English education. We constructed a desert island in Gather.town, created a story, and placed interactive objects on the desert island. In total, 25 middle school students participated in the activity for two hours. The study employed a mixed method research design, including the post-survey and interviews with the students and the teachers. The results showed that the students enjoyed the activity and responded that the activity facilitated their language learning. The interview results also confirmed that the activity was an enjoyable learning experience to them.

Conference paper

In recent years, "metaverse" has been a buzzword in our society. Big Tech companies, such as Meta and Microsoft, have been leading this trend, and its application to education is no exception (Choi, 2022). The surge of online education during the last two years also accelerated this trend. While diverse virtual reality (VR) platforms, such as Second Life and Sims, have been utilized in language-learning classrooms and their educational effectiveness has been explored in recent years, newly emerging platforms are clearly distinguished from them because these newer platforms offer distinctive features, such as proximity (e.g., users can view and communicate only with others whose avatars are close to theirs), the use of avatars, private zones (e.g., users can view and communicate only with others whose in the same private zone), video chat, and screen share, which were not possible in the previously used VR platforms (Yang & Ryu, 2021). Although new metaverse platforms in L2 learning has been rare (Fitria, 2021; McClure, 2021; Oh, 2021). The current paper examined the pedagogical possibilities of Gather.town, one of

the new metaverse platforms, in a language-learning classroom. In particular, the current study investigated how the technological affordance of Gather.town facilitated language learning. Using Gather.town, a desert island was created for Korean middle school English learners (n = 25). The content was a scenario-based, open-world adventure. The students explored the island, found interactive items, interacted with other students and teachers, completed missions, and escaped the island. The study employed a mixed method: a survey (five-point Likert scale) offered quantitative data and interviews with the students and teachers offered qualitative data.

The survey results showed that the students enjoyed the activity (mean = 4.38), particularly interaction with their peers (mean = 4.24) and the teachers (mean = 4.17) in Gather.town. The students responded that the activity helped them learn English words (mean = 4.45) and English expressions (mean = 4.38). The majority of the students wanted to participate in such activities again in the future (mean = 4.59). During the interview the students said that they liked the activity because: 1) they performed the activity with their peers, 2) they had opportunities to talk with the teachers, particularly a native English-speaking teacher, and 3) it was a new, interesting experience. They also said that they preferred the activity in Gather.town to activities in traditional classrooms or other technology-supported English learning activities. The teachers also said that they enjoyed the new experience and they found that the students quickly adopted the new environment and actively participated in the activity. However, although most of the students quickly learned how to use functions and interact with items or other people in the space, a few students struggled to understand the functions.

The current study particularly focused on contextualization and socialization (interaction between the teachers and learners and interaction among learners). For effective L2 learning, context is critical; however, unfortunately, L2 learning, particularly English as a foreign language (EFL), often lacks context (Lee & Park, 2020). Considering that language is not inseparable from the context in which it is used, a lack of context seriously undermines L2 learning effects and at the same time demotivates students (Zheng et al., 2015). The recent development of VR technologies provides an alternative space for contextualized language learning. The environment of the VR world can provide a meaningful context for L2 learning, and learners can acquire the target language in situ, which often leads to better L2 learning and retention. It can also increase learners' intrinsic motivation and interest in L2 learning (Lee, 2019; Lee & Park, 2020; Sykes & Reinhardt, 2013). Gather.town allows users to create and build their own spaces according to their purposes and needs. Using this functionality, we constructed a desert island, developed a scenario, and curated L2 learning activities on the island. The space became a learning context within which portals to different locations and interactive objects were embedded. The current study showed that this discovery learning within a context increased students' interest and motivation towards learning.

In addition, interaction was another essential key to language learning in this study. Prior studies reported several distinctive features of Gather.town, including diverse modes of communication (video, audio, and text), proximity, use of avatars, and private zones, which could benefit student learning (Fitria, 2021; Tu, 2021). The current study also found that these features particularly facilitated interaction among users. The study included diverse learner-learner and teacher-learner interactions. First, the students explored the island in pairs so that they could communicate and help each other in finding answers. Second, the students participated in a sentence-completion activity. In this activity, each pair had half of a sentence, and they found the pair who had the other half by asking around in the space. Last, the task included two interactive activities in which they interacted with the teachers by asking and answering questions. The results showed that the students preferred working in pairs and enjoyed interaction with their peers and the teachers. In fact, these interactive activities were possible due to the technological affordances of Gather.town. The proximity feature allowed the students to move around in the space and interact with each user or individual pair. Diverse modes of

communication increased their interaction among the users. In addition, the use of avatars, private zones, and proximity features enabled a safe environment, which, in turn, reduced the students' language apprehension and promoted interaction in the target language. The students also responded that the task helped them learn English vocabulary and expressions and viewed it as a valuable L2 learning experience. In sum, the results showed that the idiosyncratic features facilitated students' discovery learning, promoted interaction, and increased enjoyment of learning. This study further showed that the task in the user-created metaverse space facilitated the students' learning in the affective and language domains.

References

- Choi, S. (2022. March 4) Let's find the solution to future education in Metaverse. *The Korea Economic Daily*. https://www.hankyung.com/it/article/202203043508v
- Fitria, T. N.(2021). Creating sensation of learning in classroom: Using 'Gather town' platform video game-style for virtual classroom. *Education and Human Development Journal*, 6(2), 30-43.
- McClure, C. D., & Williams, P.N.(2021). Gather.town: An opportunity for self-paced learning in a synchronous, distance-learning environment. *Compass: Journal of Learning and Teaching*, 14(2), 1-19.
- Lee, S-M. (2019). Her story or their own stories? Digital game-based learning, student creativity, and creative writing. *ReCALL*, *31*(3), 238–254. https://doi.org/10.1017/S0958344019000028
- Lee, S-M., & Park, M. (2020). Reconceptualization of the context in language learning with a location-based AR app. Computer Assisted Language Learning, 33(8), 936–959. https://doi.org/10.1080/09588221.2019.1602545
- Oh, H. (2021). Contents application methods of metaverse platforms by type examples. *The Journal of Humanities and Social Sciences* 21, 13(1), 2673-2684.
- Sykes, J., & Reinhardt, J. (2013). Language at play: Digital games in second and foreign language teaching and learning. Upper Saddle River: Pearson.
- Tu, J. T. H.(2021). Meetings in the metaverse: Exploring online meeting spaces through meaningful interactions in Gather.town [Unpublished master's thesis, University of Waterloo]. https://uwspace.uwaterloo.ca/handle/10012/17984
- Yang, E., & Ryu, J.(2021). Effects of peer and teacher avatars on learning presence and visual attention in the metaverse learning environment. *The Journal of Educational Information and Media*, 27(4), 1629-1653.
- Zheng, D., Bischeoff, M., & Gilliland, B. (2015). Vocabulary learning in massively multiplayer online games: Context and action before words. *Education Technology Research & Development*, 63, 771–790. http://doi.org/10.1007/s11423-015-9387-4

Stephen Lambacher, Hisayo Kikuchi, Hiroyuki Obari

Aoyama Gakuin University, Tokyo, Japan

steve.lambacher@gmail.com, hisayod@gmail.com, obari119@gmail.com

Exploring the impact of AI on EFL teaching in Japan

Bio data



Stephen Lambacher is a professor in the School of Social Informatics at Aoyama Gakuin University in Japan. His research and teaching interests include ELT, CALL, and applied linguistics.



Hisayo Kikuchi is a professor in the School of Global Studies and Collaboration at Aoyama Gakuin University in Japan. She specializes in media literacy, TESOL, and applied linguistics.



Hiroyuki Obari is Professor Emeritus at Aoyama Gakuin University. He presently teaches part-time in the Faculty of Law at Waseda University and the Tokyo Institute of Technology graduate school. He specializes in worldview studies, ICT/CALL, and TESOL.

Abstract

This study investigates the use of digital personal assistants (also known as smart speakers) as part of a blended-learning (BL) environment to increase the English ability of native Japanese undergraduates. Three case studies are presented to evaluate the use of artificial intelligence (AI) smart speakers to help improve the English skills of native Japanese undergraduates by enhancing the socialization and personalization of their learning. The primary objective of the three case studies was to evaluate the efficacy of a training program that included the AI smart speakers Google Home Mini and Amazon Alexa in assisting the students in improving their English proficiency. TOEIC was utilized as a metric to ascertain if the students' English abilities improved and assess the training program's overall efficacy. Within a flipped learning (FL) environment, the case studies also incorporate 21st-century learning skills for developing international cultural awareness. Overall, mean TOEIC scores improved considerably, demonstrating that AI smart speakers helped to enhance the participants' overall English proficiency. Post-training survey results also revealed that the participants felt that using a smart speaker was a fun, easy-to-use, and practical way to improve their English speaking and vocabulary skills.

Conference paper

Introduction

AI-enabled wireless smart speakers are voice-activated devices with artificial intelligence. Voice assistants utilize a cloud-based architecture because data must be transmitted to centralized data centers. Because AI smart speakers are intended to be simple, the majority of computation and artificial intelligence processing occurs in the cloud and not on the device itself. Smart speakers are the most prevalent device with voice assistants, and they have recently begun to be used contextually in schools and institutions for advanced-level learners. The user makes a request using a voice-activated device, which is transmitted to the cloud, where the voice is converted into text. The request is then transmitted to the backend, which processes it and returns a text response. Finally, the text response is sent to the cloud, where it is converted to voice and sent back to the user (Terzopoulos & Satratzemi, 2019).

There are commands built into smart speakers that may be useful for language learners. Amazon Echo Dot, for example, can be used as a dictionary by asking "Alexa, what is the definition of...?" or "Alexa, how do you spell...?" with a given word. By asking "Alexa, what is a synonym for...?" accompanied by a word, the Echo can be used as a thesaurus or translator. There are also additional language learning skills available. Most commonly occurring are flashcard-style services, which typically link to well-known websites such as "Quizlet" or "Chegg." Moreover, Echo devices can play audio files for listening comprehension or music for relaxation or quiet study time (Davie & Hilber, 2018).

Dizon (2020) investigated whether Alexa could improve EFL listening comprehension and speaking skills. AI speakers enhanced students' speaking skills, but not their listening skills. The results of the speaking tests showed that, while the experimental group made a slight improvement, the control group did worse on the post-test than on the pretest. The author believes that improvement in L2 speaking demonstrates the potential of AI speakers to enhance foreign language development and confirms the advantages of dialogue-based CALL for language learners (see also Bibauw et al., 2019). This interaction is essential for students learning a foreign language, as they are likely to have few opportunities to speak outside the classroom.

The present study evaluates the use of digital personal assistants as part of a blended-learning environment to enhance the English language skills of native Japanese undergraduates. Three case studies are presented to ascertain the effectiveness of using AI smart speakers to improve their English proficiency and enhance the socialization and personalization of the learning. As part of a flipped learning (FL) environment, the case studies also incorporate 21st-century learning skills for fostering global cultural awareness. In each case study, pretest and post-test TOEIC scores and post-training survey results were used to evaluate the BL-training program's overall effectiveness.

Methods (Case Study 1)

The main objective of the first case was to evaluate the efficacy of a training program that incorporated the AI smart speakers Google Home Mini and Amazon Alexa to improve the English proficiency of native Japanese undergraduates. The first case study was conducted over two academic semesters (April 2018 to January 2019). Twenty-four students, all native Japanese and third-year economics majors at a private university in Tokyo, participated in the study. As part of a flipped learning environment, it incorporates 21st-century learning skills for cultivating international cultural awareness. Post-training survey results and pre- and post-test TOEIC scores were used to evaluate the overall effectiveness of the training program.

Two research questions were targeted in Case Study 1:

a. Could participants' English proficiency and understanding of 21st-century skills improve after participating in AI/BL/FL activities?

b. Did the use of AI/mobile learning and content & language integrated learning (CLIL) enhance the English skills of the native Japanese participants?

Training procedure

The technologies included Google Home Mini (Figure 1), Amazon Alexa (Figure 2), ATR CALL Brix, Facebook, Twitter, Line, and other language learning programs. TOEIC was used to determine if students' English skills improved and to evaluate the overall efficacy of the BL/FL program utilizing AI smart speakers. TOEIC was administered in April 2018 and January 2019 as a pretest and post-test, respectively.



Figure 1: Google Home mini

Figure 2: Alexa Echo Dot

Participants were divided into eight subgroups during training, with half of the groups using Google Home Mini and the other half using Amazon Alexa. Over the four-month training period, the AI smart speakers were integrated into the participants' daily lives. A timer was regularly set while interacting with the AI speakers to practice English listening, speaking, pronunciation, and vocabulary skills using a variety of software applications. For example, some groups used Google Home Mini to improve their English listening and speaking skills using the following software applications: "Best Teacher," "Travel English," "Let's play around with English," and BBC/CNN news, "Kikutan," "English Quiz" by Arc, "Liberty English," and "Kindle."

Participants recorded short video clips of their learning experiences using the AI smart speakers and uploaded them to Facebook. In addition to keeping written diaries containing their observations regarding the content and duration of their studies, participants periodically recorded their thoughts using a smartphone. After the training period, participants from each of the eight subgroups presented their impressions of the BL/FL lesson training using AI speakers. The majority indicated that the training activities positively impacted their English learning experiences.

Results (Pretest/Posttest TOEIC)

The students improved both their TOEIC listening and reading scores. Their listening mean scores increased from 185 (SD:55) to 313 (SD:76) on the pretest and post-test, respectively, and their reading mean scores also increased from 237 (SD:74) to 304 (SD:57). Pretest and post-test overall TOEIC results of Case Study 1 (n=24) during the ten-month training period (April 2018 to January 2019) increased from a mean score of 422 (SD:115) or equivalent to A1 CEFR level to 617 (SD:114) or equal to B1 CEFR level. The mean score improved by 195 points from pretest to post-test.

165

TOEIC results were analyzed using a series of t-tests, indicating the differences were statistically significant (p < .01).

Post-training questionnaire

A post-training survey was administered to 24 students to ascertain their overall impressions of the BL/FL program using AI smart speakers. Responses to a few of the survey questions are summarized as follows (n=24):

(Q1) The online lectures were beneficial in improving my English proficiency: 88% agreed.

(Q2) SNS (Facebook, Line, Twitter) helped me learn English: 92% agreed.

(Q3) This program helped me learn 21st-century skills: 90% agreed.

(Q4) Campus Crusade for Christ (CCC) members helped change my worldviews through face-to-face discussions: 88% agreed.

(Q5) Collaborative work helped improve my learning and English proficiency: 94% agreed.

(Q6) AI speaker helped improve my English skills: 76% agreed.

(Q7) AI speaker helped improve my listening skills: 87% agreed.

(Q8) AI speaker helped improve my speaking skills: 57% agreed.

(Q9) AI speaker helped improve my reading skills: 13% agreed.

(Q10) AI speaker helped improve my writing skills: 4.3% agreed.

(Q11) Presentation practice with PowerPoint helped improve my English proficiency: 100% agreed.

Method (Case Studies 2 and 3)

Two additional case studies were carried out to replicate the first case study, which revealed the benefits of using AI smart speakers in L2 learning. The second case study was conducted over two semesters (April 2019 to January 2020). Fifty-nine undergraduates participated, all native speakers of Japanese. Participants were divided into two groups: The experimental group (n=30) used a smart speaker during the training period, while the control group (n=29) did not use a smart speaker. Case Study 3 included a separate group of Japanese participants (n=23). As in Case Study 1, TOEIC was used as a measure in the additional case studies to determine if the students' English skills improved and to help ascertain the overall effectiveness of the BL/FL program. TOEIC was administered to both groups as a pretest in April 2019 and a post-test in January 2020 to determine any differences in improvement.

Training procedure

The technologies utilized in training included Amazon Alexa and Google Home Mini, ATR CALL Brix, and the mSNS programs Facebook, Twitter, and Line. During training, the experimental group participants used Amazon Alexa at home with numerous applications and filmed and kept diaries about their studies. As in Case Study 1, the AI smart speakers were integrated into their daily lives over the ten-month training period. A timer was set while interacting with the AI speakers to practice English listening, speaking, and vocabulary skills using various software programs. The control group used ATR CALL Brix online learning materials at home, focusing on listening, reading, and vocabulary learning. Therefore, the main difference between both groups was whether participants had used an AI speaker or ATR CALL Brix online materials. Pedagogical activities for both groups were nearly identical in that they studied World Heritage sites and worldviews.

A speaking test was not used in Case Studies 1 and 2. However, the participants of Case Study 3 were administered the pre-TOEIC and OPIc speaking tests in April 2019 and post-TOEIC and OPIc speaking tests in January 2020, respectively. Participants of Case

Study 3 were divided into six subgroups. Half of the participants used Google Home Mini daily to improve their English listening and speaking skills using the following software applications: "Best Teacher," "Travel English," "Let's play around with English," and BBC/CNN news. The other half of the participants used Home Mini daily to improve their listening and speaking skills using the following programs: Kikutan, English Quiz by Arc, Liberty English, and Kindle. While studying with the AI speakers, participants recorded short movie clips of their learning experiences uploaded to Facebook, just like Case Study 1. Participants also kept written diaries with their observations about the contents and duration of their studies.

Results (TOEIC and OPIc Speaking test)

Mean TOEIC scores of the experimental group improved from 407 (SD:113) to 604 (SD:92), an increase of 197 points. Mean TOEIC scores of the control group improved from 447 (SD:93) to 598 (SD:147), an increase of 147 points.

The mean TOEIC scores in Case Study 3 improved from 461 (SD: 136) to 681 (SD: 141), an increase of 229 points, and the mean OPIc speaking test scores improved from 3.9 (SD: 0.9) to 4.7 (SD: 1.25). The pre-/post-test results in both case studies were analyzed using a series of t-tests, indicating that the differences in TOEIC scores between the experimental and control groups were statistically significant (p < .01).

Post-training survey (Case Study 3)

A post-training survey was administered at the end of their respective AI/BL/FL training to ascertain their overall impressions of the program. Responses to a few questions are summarized below:

(Q1) AI speaker helped improve my English skills: 81% agreed (n=23).

(Q2) AI speaker helped improve my listening skills: 87% agreed (n=23).

(Q3) AI speaker helped improve my speaking skills: 65% agreed (n=23).

(Q4) AI speaker helped improve my reading skills: 32% agreed (n=23).

(Q5) AI speaker helped improve my writing skills: 26% agreed (n=23).

Conclusion

Mean TOEIC scores improved significantly over two semesters in each of the three case studies. The results of Case Study 2 demonstrated that the experimental group that utilized smart speakers as part of their training performed better on the TOEIC post-test than the control group that did not use smart speakers. The results of Case Study 3 also revealed that the integration of AI smart speakers and flipped learning helped the experimental participants improve their TOEIC scores by 229 points. The follow-up survey revealed that over 80% of participants agreed that the smart speaker improve their English skills, particularly their listening skills. The participants also felt that using a smart speaker was a fun, easy-to-use, and practical way to improve their English speaking and vocabulary skills. The smart speaker's ability to comprehend their commands instilled confidence in the participants, and they also enjoyed being exposed to authentic native pronunciation by listening to the news.

AI smart speakers appear helpful for developing speaking skills, particularly pronunciation (e.g., Dizon, 2020; Obari et al., 2020). Students can experience valuable listening and speaking practice outside of the classroom with the help of AI speakers, resulting in greater learner autonomy (Dizon & Tang, 2020). The findings of our three case studies support this assertion. Although interaction with a native speaker is preferable, some students may feel uneasy speaking the target language in front of others or have limited opportunities to converse with L2 speakers outside of formal settings. Consequently, an AI smart speaker could serve as an excellent

language-learning tool. Although research on the application of smart speakers for L2 learning is still in its infancy, further studies with more subjects and varied L2 learning contexts would be helpful. Another area of inquiry might focus on learner and teacher training in utilizing smart speakers more effectively.

References

- Bibauw, S., François, T., & Desmet, P. (2019). Discussing with a computer to practice a foreign language: Research synthesis and conceptual framework of dialogue-based CALL. *Computer Assisted Language Learning*, 32(8), 827–877. https://doi.org/10.1080/09588221.2018.1535508Davie, N., & Hilber, T. (2018). Opportunities and challenges of using Amazon Echo in education. *International Association for Development of the Information Society*.
- Dizon, G. (2020). Evaluating intelligent personal assistants for L2 listening and speaking development. Language Learning & Technology, 24(1), 16–26. https://doi.org/10.1177/00336882211020548
- Dizon, G., & Tang, D. (2020). Intelligent personal assistants for autonomous second language learning: An investigation of Alexa. *The JALT CALL Journal*, 16(2), 107–120. https://doi.org/10.29140/jaltcall.v16n2.273 https://doi.org/10.29140/jaltcall.v16n2.273
- Obari, H., Lambacher, S., & Kikuchi, H. (2020, June). Utilizing AI smart speakers to improve the English skills of Japanese university students. Paper presented at the JALTCALL 2021 Conference: Remote Teaching and Beyond, Aomori, Japan.
- Terzopoulos, G., & Satratzemi, M. (2019, September). Voice assistants and artificial intelligence in education. In G. Eleftherakis, M. Lazarova, A. Aleksieva-Petrova,& A. Tasheva (Eds.), *Proceedings of the 9th Balkan Conference on Informatics* (pp. 1–6). Association for Computing Machinery. https://doi.org/10.1145/3351556.3351588

Veronica Juliana Schmalz, Frederik Cornillie

ITEC, imec research group at KU Leuven, Kortrijk, Belgium

veronicajuliana.schmalz@kuleuven.be, frederik.cornillie@kuleuven.be

Towards truly intelligent and personalized ICALL systems using Fluid Construction Grammar

Bio data



Veronica Juliana Schmalz is a PhD student working on the development of a computational model for the usage-based acquisition of modular constructions and grammatical categories. Apart from computational linguistics and AI, she is also interested in language acquisition and development, multilingualism, cognitive linguistics and assessment of learners' competences.



Frederik Cornillie (Ph.D., KU Leuven) is research and valorization manager in educational technology at KU Leuven and at the strategic research institute imec in Belgium. His main research interest in CALL is the intersection of tutorial CALL and task-based language learning, and more specifically how mechanics of games such as feedback can support instructed language learning. He is also passionate about the design and development of CALL applications.

Abstract

Intelligent Computer-Assisted Language Learning (ICALL) aims to design effective systems for the analysis of learners' production in a target language ensuring both successful learning and motivated learners. Most of the existing systems, however, focus extensively on the form rather than on the meaning of language. To create effective systems facilitating personalized language learning both form and meaning should be considered. The reason behind this is that language is a continuous flow of information passing from one user or agent to another, both during comprehension and production. This becomes even more relevant in the case of second or foreign languages (L2), where certain linguistic choices may be dictated by inexact form-meaning links construed by the learner. In this research project, we focus on the analysis of the spoken production of adult learners of German, taking argument and information structure as a use case. We use Fluid Construction Grammar as a formalism, which captures relevant linguistic aspects at both the syntactic level (form) and the semantic level (meaning). Its particularity lies in the possibility of closely monitoring bidirectional form-meaning interactions starting from constructions of different nature modeled in an extensively customizable way. Our work is in progress, and we focus on ways to provide helpful feedback on meaning. German displays a rather articulated grammar and obtaining insights not only on its formal but also on its semantic correctness could offer important steps forward for Intelligent CALL.

The design of computational systems for Intelligent CALL that can effectively support L2 learners in personalized learning requires a grammatical framework that is computationally effective and offers linguistic and acquisitional perspicuity (Schulze & Penner, 2008).

Proceedings of the XXIst International CALL Research Conference

Since human language strictly depends on users and on contexts, we need a system that:

- models the learners' productions in a highly detailed and context-specific way;
- bidirectionally and simultaneously accesses the form and the meaning of the learners' productions highlighting their intersections;
- compares several productions focusing on different language units, also considering their frequency and logical plausibility.

These desiderata are currently found in Fluid Construction Grammar (FCG) (Steels, 2017), a formalism that originated from the intersection of linguistics and AI methods. FCG exploits pattern-finding, generalization and specialization principles (Van Eecke, 2018) to model grammars in the form of constructions, or bidirectional form-meaning pairs. FCG can formally represent elementary linguistic structures, such as lexical units or nominal phrases, and more complex ones, such as idioms. Its application with a free and user-friendly editor (https://www.fcg-net.org/download/) on datasets like ours, containing L2 learners' productions (Baten & Cornillie, 2019), allows us to eventually reach a level of personalization in intelligent tutorial CALL as we have not seen so far. This is due to the possibility of engineering customizable constructions from scratch and concentrating on specific aspects, such as the German argument and information structure, along with the case system (van Trijp, 2011) for our current investigation. Moreover, constructions can have added features that allow their generalization to specific productions, and even the detection of formal or semantic errors. Thanks to FCG and the close observation of form-meaning aspects, we can tackle the following challenges related to either the psychological or the ecological dimension of Smart CALL as defined in the conference theme:

- provide learners with personalized feedback going beyond formal inaccuracies so as to improve their effort/reward ratio and mental acceptance of tutorial CALL activities;
- offer additional information to the teachers or the learners' more skilled peers derived from the application of linguistic constructions, creating opportunities for 'distributed scaffolding' (Tabak & Kyza, 2018) co-provided by the teacher and the CALL system in the ecology around the learner;
- deliver linguistic data to the research community for easily accessible semantically annotated corpora, fundamental to the creation of truly intelligent tutors who can efficiently communicate with learners and provide them with the necessary personalized learning inputs and feedback (Beuls, 2013).

Conference paper

Introduction

Since its establishment, one of the principal goals of Intelligent Computer-Assisted Language Learning (ICALL) was the design of effective systems capable of identifying and correcting the errors of language learners, as well as accurately encoding their competence (Melissa et al., 1993). To measure the effectiveness of such systems, it is crucial to take into account their capacity to enhance the learning experience and improve the motivation of learners. Nevertheless, most existing systems mainly focus on the form rather than on the meaning of language (cf. Schulze & Penner, 2008). In order to take a significant step forward towards the development of more efficient and personalized ICALL systems, the correlation between form and meaning in grammar should receive more attention. The reason behind this is the bidirectionality between the two parts, which turn out to be continuously correlated both in language comprehension and formulation (Goldberg, 2003). This appears especially crucial when learning a new language, since learners need to focus on properly understanding and correctly generating utterances in a different language than their native one.

The acquisition of a new language system, namely of its grammar, represents a demanding challenge since each language has a unique way of representing reality through distinct syntactic and semantic units and principles. Monitoring the interactions between these different components in a clear and understandable way proves to be a complex task as well, since they are thoroughly intertwined and subject to variation. Construction grammars come in response to this daunting task. Their name refers to their fundamental element, that is the construction, a more or less complex unit consisting of a single or multiple words with different functions, characterized by a form and a meaning, as well as other syntactic and semantic features attributed to it (Goldberg, 2006). Construction grammars have been previously regarded as a suitable framework for ICALL systems by Schulze and Penner (2008) since they meet the requirements of *computational effectiveness, linguistic perspicuity* and *acquisitional perspicuity* introduced by Matthews (1997).

According to usage-based theories, mainly based on general cognitive capacities such as intention reading and pattern finding (Tomasello, 2003), construction grammars can be used to faithfully model the processes of first-language acquisition. Through the years several formalisms have emerged, among which:

- HPSG Head-driven Phrase Structure Grammar (cf. Copestake & Flickinger, 2000)
- ECG Embodied Construction Grammar (cf. Bergen & Chang, 2005)
- SBCG Sign-Based Construction Grammar (cf. Michaelis, 2009)
- FCG Fluid Construction Grammar (cf. Steels, 2011)

Recent computational implementations have made it possible to investigate in more detail the adequacy of construction grammars to model foreign language learning strategies. Among these, FCG constitutes the only bidirectional framework to computationally implement construction grammars. It originated in the 1990s from the combination of experts' knowledge in the artificial intelligence domain, as well as the cognitive and formal linguistics domain. Initially, it was adopted to model the process of grammar acquisition in artificial agents (Steels at al., 2012). Today, it can be applied to more general or specific domains, such as representing and monitoring the usage-based acquisition of a language's grammar (Doumen et al., 2021) or investigating the case system in German (van Trijp, 2011). Particularly, it enables the mapping between an utterance and its meaning, as well as the opposite from the meaning, which can be expressed in any form, to the utterance. This can be all managed through a user-friendly interface¹ that displays the different engineered constructions individually and their progressive interactions. The uniqueness and strength of FCG lies in its *fluidity*, namely in the possibility of endless customization and the addition of linguistic features. This provides a robust way to parse and analyze in detail linguistic data like ours, as well as to design target corrections and feedback for learners (see Sections 5 and 6).

Case study: Argument and information structure in German

In this research project, we focus on the analysis of 1,487 transcribed and annotated spoken productions of 36 students of German as a second language (L2) from the academic program in Languages and Literature at the University of Ghent, collected in an oral elicited imitation task (Baten & Cornillie, 2019). The learners heard an audio stimulus containing transitive and ditransitive verbs, along with noun and prepositional phrases, which they had to match to one of two displayed pictures with different semantic roles. Subsequently, they needed to produce an oral response describing the selected picture, which should have corresponded to the initially heard utterance's meaning (see Figure 1).

¹ https://www.fcg-net.org/

Proceedings of the XXIst International CALL Research Conference


Figure 1. Two pictures containing different semantic representations of an event are presented to the students. In the left one the mother goes to the shop <u>with</u> the son, while in the right one she does it <u>without</u> the son. The stimulus received by the students concerns the picture on the right. In the example response the student produces a semantically correct utterance, which however contains a formal error in case selection. The preposition <u>'ohne'</u> requires an accusative and not a dative determiner.

We are interested in the acquisition of the German argument and information structure. We explore how these can be modeled based on the case system and Abstract Meaning Representation (AMR) rules for semantic roles (cf. Banarescu et al., 2013) in a tailor-made FCG grammar. Exploiting the versatility of FCG and a user-friendly interface, our aim is first to create a grammar capable of a bidirectional representation of the stimuli, namely 48 utterances used to describe the pictures. Once this grammar has been created and tested both in comprehension-namely providing an utterance FCG outputs its meaning (including the argument structure with semantic roles) and in formulation—namely providing an AMR meaning FCG can generate a desired utterance, we can proceed to model the learners' responses. To do so, we, as grammar engineers, have to implement a series of rules that apply in the case of correct and incorrect responses. In the case of the latter, there are two different approaches to handle students' errors: mal-rules and constraint relaxation. The former consists in the creation of rules based on learners' errors using mal-rules (cf. Sleeman, 1982; Matthews, 1992). The latter relaxes the constraints of the original grammatical rules so that they can also apply in the case of errors (Foth et al., 2005). Since there is no set of rules capable of capturing all the errors expected by learners, the two aforementioned methods can be combined in FCG to detect errors and provide feedback, as well as to allow the processing and understanding of learners' productions (see Sections 5 and 6).

Fluid Construction Grammar (FCG)

Differently from other computational construction grammars, FCG does not have specific implementation principles. Therefore, it is suitable for representing how learners, or in general language users, perceive and utter about the world using grammar. Importantly, FCG relies on the principle of bidirectionality between form and meaning. As a consequence, we can exploit it both to comprehend given utterances and to formulate new ones. Each FCG construction, namely the linguistic unit, needs to have a semantic part related to meaning and a syntactic part related to form.

application process		1, 2.10: mutter-cxn (exn 0.50)		
		succeeded, cxn-appl	succeeded, cxn-applied		
		status	cxn-applied		
		source structure	transient structure		
			* root		
		applied construction	mutter-cxn (cxn 0.50) show attributes		
			?mother-word		
			referent: ?m		
			syn-cat: ?mother-word		
	0, 1.00: initial		ex-class: noun		
			Case: ((ri) - /i) - /) (rai - rai - /) (rgi - /gi - /) (rdi - /(i - /) (+ + + -))		
			animacy; animate		
		resulting structure	transient structure		
			mutter-6		
		resulting bindings	((?mother-word . mutter-6))		
		meaning	(mother 7m-2004)		
			Second Change		
constructional dependencies					
	matter-exp				
	'trother-word				
applied constructions	mutter-cxn (ca	n 0.50)			
resulting structure	transient structure				
	root				
	and the second				
	meaning: (mother(7m-2806)) form: (strinorimuter-6, "mutter"))				
	⊙ syn-cat: lex-class: noun				
	case: ((?)	if-1800 - ?nf-1800), (?	af-2324 - ?af-2324), (?gf-1011 - ?gf-1011), (?df-2437 - ?df-2437), (+ - +)]		
	reterent: ?	m-2906			
	animacy:	animate			

Figure 2. Lexical construction 'Mutter' with form-meaning information displayed in the application process of constructions and resulting transient structure

There can be different types of constructions, such as lexical constructions for nouns with a basic meaning and form, along with some basic semantic (e.g. *animacy*) and syntactic (e.g., *lex-class noun*) information (see Figure 2), and phrasal ones that combine lexical constructions for nouns with other units, for example determiners or prepositions (see Figure 3).

application process	0, 1.00: Initial 4 .4.50: contracted-prep-phrase-cxn (cxn 0.50) 4.4.50: contracted-prep-phrase-cxn (cxn 0.50)
constructional dependencies	
	auen-xxn contracted-prep-phrase-xxn 7atore-word 7aoun
	Zum-cxn ?contracted-prep
	?to-word ?contracted-prep-phrase
applied constructions	zum-exn (exn 0.50)
	laden-cxn (cxn 0.50)
	contracted-prep-phrase-cxn (cxn 0.50)
resulting structure	transient structure
	root
	contracted-prep-phrase-6 iden-1

Figure 3. Comprehension of the prepositional phrase 'zum Laden'. The application process represents how smaller constructions contribute to a larger prepositional phrase construction, based on a dative masculine contracted preposition 'zum' and a masculine noun 'Laden'

Proceedings of the XXIst International CALL Research Conference

Concomitantly, constructions can map meaning to semantic categories but also syntactic categories to a form, namely a word or phrase, and semantic categories to syntactic units. To represent how constructions are used by a language user or learner, they are arranged in a data structure called *transient structure*. It displays how the processing of a linguistic utterance happens during the application of constructions.



Figure 4. Resulting transient structure and meaning networks for the comprehension of the utterance 'die Mutter geht ohne den Sohn zum Laden'

Constructions and transient structures regularly interact with each other via matching and merging processes. For instance, in comprehension an initial transient structure, namely a visual representation of how the form is organized, is built. Afterwards, the matching checks whether the features of existing constructions can apply to the given stimuli. Therefore, if successful, the transient structure receives additional information. The final transient structure is obtained after all constructions have applied. It contains morphological and lexical features as well as data relative to the information and argument structure (see Figure 4).

A computational model for the German grammar

In the last years, FCG has been used to implement different grammars focusing on distinct aspects, for example the Spanish verb conjugation (Beuls, 2013), English quantifiers (Pauw & Hilferty, 2012), Russian verbal aspect (Gerasymova, 2012), Hungarian verbal agreement (Beuls, 2011), Polish negation (Höfer, 2012), L2 Spanish verb morphology (Beuls, 2014), German spatial language (Spranger & Loetzsch, 2011) and German case system (van Trijp, 2011). Since FCG is a dynamic, adaptive and flexible tool, it can be efficiently exploited to represent how languages are continuously influenced and modified by language users' diverse usages and variations. This makes it a perfectly functioning formalism for our project since we are handling learners' productions in a rich and elaborate grammatical system like the German one.

One of the distinctive aspects of German is its rich and articulated case system. It can be a rather complex concept to learn, especially for those learners whose native languages do not exhibit such a range of case options. The information indicating the case is often encapsulated in the determiner or in the suffixes of adjectives, and in some cases also in nouns. However, the presence of syncretism, that is the polyvalence of determiners or suffixes for case, number and gender, can often cause issues in ruling out one single case among the different options and assigning a syntactic role to a given construction. For example, a sentence containing two determined feminine nouns as subject and as object can create confusion, especially when the argument and information structure do not follow the standard order and one is topicalized (e.g., <u>die Katze object sieht die Gans subject</u> aus der Ferne – the cat sees the goose from far away).

Moreover, when dealing with prepositions that require a specific case to acquire a certain meaning, the correct application and recognition of the case system is crucial. These aspects of German grammar are present in our students' data. Therefore, our task was to model them using FCG in the most efficient and least ambiguous way possible. To solve the case syncretism issue in computationally modeling German grammar, we decided to follow the features with binary values used in phonology and based on the application of unification processes. Despite its apparent simplicity, it allows to combine information related to gender, number and case, and compares it across multiple levels and units. In a similar fashion, the matrix is used for the argument structure of verbs. For example, transitive verbs need the nominative case for the agent, or subject, and the accusative for the patient, or object (see Figure 4 - meaning network). The task is made more complex in the case of prepositional phrases with motion verbs and topicalized arguments.

Each verb presents a different argument structure depending on the macro-category to which it belongs according to the formalism of AMR (cf. Banarescu et al., 2013). AMR is a semantic representation that makes use of PropBank framesets (Kingsbury & Palmer 2002) and labeled graphs to analyze and connect different parts of speech in an understandable way both for human users and artificial agents. The goal of AMR is to be an intuitive and easily interpretable, yet consistent and inclusive annotation. Since we want our grammar to be applicable both in formulation and in comprehension, we also added features to rule out the possible alternatives among several prepositional phrases and information structures. In fact, we distinguish contracted and non-contracted prepositional phrases, but also stative, motion, means and accompaniment prepositional phrases which can match specific verbs arguments. On the other hand, for what concerns the information structure, we exploit FCG possibilities to model it separately from the argument structure, highlighting variations from the non-topicalized standard order.

	56, 1.00: intransitive-extra-argument-structure-cxn (cxn 0.50) second-merge-failed	1	
	55, 14.30: incorrect-loc-intransitive-extra-argument-structure	-cxn (cxn 0.50)	
ľ	transient structure		
	root		
		accompanying-phrase-poss-396 ihren-1 ohne-4	57, 15.80: topic-arg0-extra-info-arg4-information-structure-cxn (cxn 0.50)
	incorrect-loc-intransitive-extra-argument- structure-unit-2	incorrect-contracted-prep- phrase-1131	58, 1.00: Intransitive-extra-argument-structure-cxn (cxn 0.50) second-merge-failed
		noun-phrase-1209	

Figure 5. Constructions' application process in the comprehension of an incorrect intransitive utterance. Red constructions could not be applied, while green ones could. The error here is caused by the undetermined prepositional phrase 'zu Laden' which inhibits the application of the argument structure construction but not of the independent information structure

Detecting errors

176

We have engineered an FCG grammar according to a hand-written set of specific constructions related to the stimuli sentences that were provided to the students during the oral elicited imitation task. With this grammar, correct German sentences can be comprehended and automatically generated. We can extend this by adding the variations produced by the learners. These can be of three types:

- formal errors denoting lack of competence in certain grammatical aspects,
- <u>meaning errors</u> signaling incorrect understanding of the stimuli or interpretation of the argument structure,
- <u>variations</u> with respect to the received stimulus but still formally and semantically correct.

When attempting to parse a student response containing one or more of the above variations, the created computational model of German grammar detects differences from the norm. These differences are identified and visually displayed in the FCG web interface (see Figure 5).

For the variations found in our data, contained in 282 sentences, our approach is to implement mal-rules that allow us to go beyond error detection and continue with the comprehension of the students' sentences using our grammar. These are part of a specific construction set that comes into play whenever the pre-existing constructions of the designed grammar fail to apply (see Figure 6). We can also add information containing clarifications about the variations, such as indications on what we would have expected and why these are considered incorrect or improper. This constitutes the basis for providing feedback on meaning, rather than exclusively on form (see Section 6).



Figure 6. *Example of mal-rules (in orange) in the application process for the comprehension of `die Mutter geht ohne dem Sohn zum Laden'. In this error case, the dative instead of accusative case is used with the preposition `ohne'.*

Providing meaningful feedback

Given its bidirectionality and open-endedness, FCG allows us to handle form and meaning errors, as well as variations in detail, apart from simply detecting them. These two possibilities prove its adequacy for the analysis of language learners' production as well as for the design of didactically relevant feedback, tailored to the specifically targeted language aspects and learners' metalinguistic knowledge. For example, when a student utters a sentence such as 'der Doktor verkauft den Clown das Buch', although each noun phrase has a correctly matched determiner, its meaning is incorrect since the AMR required arguments for the ditransitive verb 'verkaufen – to sell' are *arg0* (subject-seller) in nominative, *arg1* (object-sold entity) in accusative and *arg2* (receiver-buyer) in dative. This error is captured by a specific mal-rule that we complemented with additional information concerning the error type and the reason for it. The consequences of the error can also be seen in the resulting transient structure where the arg2 field is empty and there are two possible arg1s (see Figure 7).



Figure 7. Incorrect construction with mal-rule application and meaning network for the utterance 'der Doktor verkauft den Clown das Buch' with two accusatives instead of a dative case for arg2.

Proceedings of the XXIst International CALL Research Conference

In addition to this type of cases, we have cases in which the learner modifies the stimulus sentence more freely. For example, instead of 'die Mutter geht ohne den Sohn zum Laden' a learner can say 'die Mutter geht ohne ihren Sohn zum Laden'. Although it deviates from the expected answer, considering the original stimulus, it is not ungrammatical, and the meaning does not differ from the expected response. Therefore, it can be accepted as a variation, and the grammatical constructions can apply effortlessly, as long as the new word used is added to the lexicon.

With these methods we are trying to go beyond error detection and personalize the feedback that learners receive depending on their metalinguistic awareness and competence, as well as on the areas of concern of a given exercise or test in a foreign language.

Conclusion

With this project we are experimenting with the use of a computational construction grammar formalism to represent not only the engineering of the grammar of a correct and established language but also variations on the target language as produced by language learners. Exploiting the possibility of extensive customization, the modularity of constructions and the efficiency of the FCG editor in visualizing the interactions between form and meaning, as well as between argument and information structure, FCG appears to be a promising tool to be used in the study of language acquisition processes, even beyond its original conception for artificial agents. The ability to precisely detect the source and site of errors in the path of application of constructions brings us closer to the reasoning that possibly a learner makes and enables us to provide more in-depth and personalized feedback. In order to do that, we are closely analyzing our dataset, especially the students' responses, searching for generalizable patterns and possibilities to scale up the design and application of constructions. However, this is an uphill climb that requires close collaboration between researchers in computational linguists and language acquisition, teachers, and the learners themselves in order to bring significant improvements and new insights to ICALL.

References

- Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., Knight, K., Koehn, P., Palmer, M., & Schneider, N. (2013). Abstract meaning representation for SemBanking. In Proceedings of the 7th linguistic annotation workshop and interoperability with discourse, 178-186.
- Baten, K., & Cornillie, F. (2019). Elicited imitation as a window into developmental stages. Journal of the European Second Language Association, 3(1), 23–34.
- Bergen, B. K., & Chang, N. (2005). Embodied construction grammar in simulation- based language understanding. In J-O. Östman & M. Fried (Eds.), *Construction* grammar(s): Cognitive and cross-language dimensions (pp. 147-190). Amsterdam: John Benjamins.
- Beuls, K. (2011). Construction sets and unmarked forms: a case study for Hungarian verbal agreement. In L. Steels (Ed.), *Design patterns in fluid construction* grammar (pp. 237–264). Amsterdam: John Benjamins.
- Beuls, K. (2013). Towards an agent-based tutoring system for Spanish verb conjugation. [Doctoral dissertation, Vrije Universiteit Brussel].
- Beuls, K. (2014) Grammatical error diagnosis in fluid construction grammar: A case study in L2 Spanish verb morphology, *Computer Assisted Language Learning*, 27(3), 246-260.
- Copestake, A., & Flickinger, D. (2000). An open-source grammar development environment and broad-coverage English grammar using HPSG. In Proceedings of the 2nd international conference on language resources and evaluation (LREC 2000), 591–600.

Croft, W. (2001). Radical construction grammar. Syntactic theory in typological perspective. Oxford: Oxford University Press.

Doumen, J., Beuls, K. & Van Eecke, P. (2021). A computational model of the co-emergence of item-based constructions and syntactic categories. 11th International Conference on Construction Grammar (ICCG11).

Foth, K., Menzel, W., & Schröder, I. (2005). Robust parsing with weighted constraints. *Natural Language Engineering*, 11(1), 1–25.

Gerasymova, K. (2012). Expressing grammatical meaning with morphology: A case study for Russian aspect. In L. Steels (Ed.), *Computational issues in fluid construction* grammar (pp.91-112). Berlin: Springer Verlag.

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5), 219-224.

Goldberg, A. E. (2006). Constructions at work: The nature of generalization in language. Oxford University Press on Demand.

Höfer, S. (2012). Complex declension systems and morphology in Fluid Construction Grammar: A case study of Polish. In L. Steels (Ed.), *Computational issues in Fluid Construction Grammar* (pp.143-177). Berlin: Springer Verlag.

Kingsbury, P.R. & Palmer, M. (2002). From TreeBank to PropBank. LREC.

Matthews, C. (1992). Going AI: foundations of ICALL. *Computer Assisted Language Learning*, *5*(1-2), 13–31.

Matthews, C. (1997). Grammar frameworks in intelligent CALL. CALICO Journal, 11(1). 5–27.

Melissa, H. V., Maisano, R., Alderks, C. & Martin, J. (1993). Parsers in tutors: What are they good for?. CALICO Journal, 11(1), 28–46.

Michaelis, L. A. (2009). Sign-based construction grammar. In B. Heine & H. Narrog (Eds.), *The Oxford handbook of linguistic analysis* (pp. 155-176), Oxford: Oxford University Press.

Pauw, S., & Hilferty, J. (2012). The emergence of quantifiers. Experiments in cultural language evolution, 3, 277-283.

Schulze, M., & Penner, N. (2008). Construction grammar in ICALL. Computer Assisted Language Learning, 21(5), 427–440.

Sleeman, D. (1982). Inferring (mal) rules from pupils' protocols. Selected and updated papers from the proceedings of the 1982 European conference on Progress in artificial intelligence, 30-39.

Spranger, M., & Loetzsch, M. (2011). "Syntactic indeterminacy and semantic ambiguity." In L. Steels (Ed.), *Design patterns in fluid construction grammar* (pp. 237–264). Amsterdam: John Benjamins.

Steels, L. (2011). (Ed.). *Design patterns in Fluid Construction Grammar*. Amsterdam: John Benjamins.

Steels, L., De Beule, J., & Wellens, P. (2012). Fluid construction grammar on real robots. In L. Steels & M. Hild (Eds.), *Language grounding in robots* (pp. 195–213). Berlin, Heidelberg: Springer.

Steels, L. (2017). Basics of fluid construction grammar. *Constructions and Frames*, 9(2), 178–225.

 Tabak, I., & Kyza, E. A. (2018). Research on scaffolding in the learning sciences. International Handbook of the Learning Sciences (pp. 191–200). New York: Routledge. Updated papers from the proceedings of the 1982 European conference in artificial intelligence (pp. 30–39). John Wiley & Sons, Inc.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Van Eecke, P. (2018). Generalisation and specialisation operators for computational construction grammar and their application in evolutionary linguistics research. [Doctoral dissertation, Vrije Universiteit Brussel].

van Trijp Remi (2011). Feature matrices and agreement: A case study for German case. In Luc Steels (Ed.), *Design patterns in fluid construction grammar* (pp. 205–235). Amsterdam: John Benjamins.

Proceedings of the XXIst International CALL Research Conference

Ward Peeters

Kanda University of International Studies, Chiba, Japan

peeters-w@kanda.kuis.ac.jp

Computer-mediated communication (CMC) text analytics: Exploring the dynamics within digital discourse

Bio data

Ward Peeters holds a PhD in Linguistics from the University of Antwerp (Belgium). He is a research fellow in science communication at Leiden University (the Netherlands) and a visiting professor at Kanda University of International Studies (Japan). His main interests revolve around studying social network impact in language learning contexts, which have materialised in a number of research projects in Belgium, Japan, the Netherlands and South Africa as part of an extensive study on computer-supported collaborative work.

Abstract

While the scholarship on computer-mediated communication (CMC) can be described as a diverse field of study, with a variety of niche areas and transdisciplinary domains, one of the main recurring research practices within the field involves analytic examinations of digital discourse (Vásquez, 2022). These examinations generally include qualitative and quantitative measures of digital text which is produced, used and re-used in a fast-evolving communicative landscape. Because of the rapid succession of devices, programs and applications that are used for mediation, the methods for analysis sometimes lag behind (as discussed by Herring, 2019). This has raised the question among scholars on how to effectively study and visualise digital discourse, taking into account all affordances and contextual factors that digital resources provide (Car, 2020; Lin, 2015). This paper presents a number of key issues and challenges in the field, bringing together findings from Computer Assisted Language Learning and Learning Analytics studies to provide some new perspectives on CMC text analytics for education.

Conference paper

Introduction

This paper discusses some of the new scholarship on computer-mediated communication (CMC) text analytics, drawing on findings from Peeters (2022), published in Smart CALL: Personalization, contextualization, & socialization (Colpaert & Stockwell, 2022), a volume which brings together new theoretical perspectives on technology and language learning. In this chapter, the application of social network analysis (SNA) principles—which have proven to generate substantial insights into the dynamics of mediated digital text—has been elaborated upon. Special attention was paid to time measures in this context, showing how interaction and collaboration patterns can change over time and how relevant time frames should be determined to inform any analysis of digital discourse.

In addition to the findings of Peeters (2022), this paper briefly discusses the use of epistemic network analysis (ENA) principles to analyse digital discourse. ENA, which can be defined in brief as an approach to analyse the connections and co-occurrences of certain 'codes' in dialogue, can be used to generate dynamic network models out of

digital text (Gašević et al., 2019; Peeters et al., 2020; Shaffer, 2017). ENA has been found to provide additional insights into the dynamics of CMC, as well as into associated learning and teaching processes (Vandenberg, 2021; Vujovic et al., 2021).

In the next section, the application of SNA and ENA for digital text analytics will be briefly discussed.

Network Analyses and digital discourse

Applications of social network analysis (SNA)

The chapter by Peeters (2022) discusses how the application of SNA principles to analyse and visualise the process of socialisation and information exchange in CMC text allows us a look into the wider semiotic contexts in which the learning process is taking place. In the presented case, we are able to see how socialisation unfolds while students collaborate on a number of learning tasks through Google Classroom. SNA principles were also applied to visualise the interplay between different aspects of socialisation over time.

Peeters (2022) argues that it is both informative and educational for educators and researchers to be able to observe the shifts that can occur in such interactive spaces for learning. This way we can improve both the design of CMC spaces for education—as we can create an overview of the dynamics that are present or absent in a certain learning space— as well as the support mechanisms that educators can provide throughout the learning process of their students.

Applications of Epistemic network analysis (ENA)

As mentioned in the introduction, ENA is a quantitative ethnographic approach to analyse patterns and dynamics in coded dialogue and has been designed to compare the weighted structure of links between different coded elements in discourse (Gašević et al., 2019; Peeters et al., 2020; Shaffer, 2017). It provides researchers with methods to empirically determine the behaviour and function of different discourse elements, and to build networks in which they can examine, among others, the co-occurrence or co-existence of different topics, events, emotions or collaborative actions.

The principles of ENA, together with other process measures, have been applied to online learner data before, including individuals' use of learning strategies, community of inquiry frames and self-regulated learning techniques (e.g., Bressler et al., 2019; Saqr et al., 2021). Because the analytic technique "was originally designed to capture learning as it is expressed in conversation turns, and in collaborative scenarios" (Saint et al., 2021, p. 10), it is no surprise that it has been deemed a powerful tool to analyse collaborative digital discourse as well.

Oshima and Shaffer (2021) describe in their paper on learning analytics for a new epistemological perspective of learning that ENA enables researchers to easily create visual comparisons between selected codes and variables. Referring to an earlier publication (Shaffer & Ruis, 2017), they also explain that most network analyses and accompanying statistics tend to describe how nodes or actors are connected, but gloss over the content, or what these nodes represent.

In combination with probabilistic and temporal measurements, ENA enables researchers to look at the way codes co-exist over time, how influential they are in the network, how relationships evolve and how consequential they are in the network's evolution.

Conclusion

As detailed in Peeters (2022), the principles of SNA can be applied to generate new insights into the dynamics of mediated digital text and the structure of a discourse network. ENA principles, on the other hand, can shed more light on the contents of mediated text and how certain actions, strategies or learning paths co-occur and co-exist. By taking into account relevant time intervals, both SNA and ENA can be used to visualise how these dynamics evolve over time, providing us with the necessary methods and insights to determine when and how learners communicate and collaborate.

References

182

Bressler, D. M., Bodzin, A. M., Eagan, B. & Tabatabai, S. (2019). Using epistemic network
analysis to examine discourse and scientific practice during a collaborative game.
Journal of Science Education and Technology, 28(5), 553–566.
https://doi.org/10.1007/s10956-019-09786-8
Carr C T (2020) CMC is dead, long live CMCL: Situating computer-mediated

- Carr, C. T. (2020). CMC is dead, long live CMC!: Situating computer-mediated communication scholarship beyond the digital age. *Journal of Computer-Mediated Communication*, 25(1), 9–22. https://doi.org/10.1093/jcmc/zmz018
- Colpaert, J., & Stockwell, G. (Eds.), *Smart CALL: Personalization, contextualization, & socialization*. London: Castledown Publishers. https://doi.org/10.29140/9781914291012
- Gašević, D., Joksimović, S., Eagan, B., & Shaffer, D. W. (2019). SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, 92, 562–577. https://doi.org/10.1016/j.chb.2018.07.003
- Herring, S.C. (2019). The coevolution of computer-mediated communication and computer-mediated discourse analysis. In P. Bou-Franch & P. G. C. Blitvich (Eds.), Analyzing digital discourse (pp. 25–67). Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-319- 92663-6_2
- Lin, H. (2015). A meta-synthesis of empirical research on the effectiveness of computer-mediated communication (CMC) in SLA. *Language Learning & Technology*, 19(2), 85–117. http://doi.org/10125/44419
- Oshima, J., & Shaffer, D. W. (2021). Learning analytics for a new epistemological perspective of learning. *Information and Technology in Education and Learning*, 1(1), Inv-p003. https://doi.org/10.12937/itel.1.1.Inv.p003
- Peeters, W., Saqr, M., & Viberg, O. (2020). Applying learning analytics to map students' self-regulated learning tactics in an academic writing course. In H. J. So, M. M. Rodrigo, J. Mason, & A. Mitrovic (Eds.), *Proceedings of the 28th ICCE* (pp. 245–254). Taiwan: APSCE. https://www.self.com/academic/aca

https://apsce.net/icce/icce2020/proceedings/paper_143.pdf

- Peeters, W. (2022). New perspectives on computer-mediated communication research: A social network analysis approach. In J. Colpaert, & G. Stockwell (Eds.), Smart CALL: Personalization, contextualization, & socialization (pp. 29–54). London: Castledown Publishers. https://doi.org/10.29140/9781914291012-3
- Saint, J., Fan, Y., Gašević, D., & Pardo, A. (2022). Temporally-focused analytics of self-regulated learning: A systematic review of literature. *Computers and Education: Artificial Intelligence*, 100060. https://doi.org/10.1016/j.caeai.2022.100060
- Saqr, M., Peeters, W., & Viberg, O. (2021). The relational, co-temporal, contemporaneous, and longitudinal dynamics of self-regulation for academic writing. *Research and Practice in Technology Enhanced Learning*, 16(1), 1–22. https://doi.org/10.1186/s41039-021-00175-7
- Shaffer, D. W., & Ruis, A. (2017). Epistemic network analysis: A worked example of theory-based learning analytics. In C. Lang, G. Siemens, A. Wise, & D. Gasevic,

(Eds.), *Handbook of learning analytics* (pp.175–187). New York: SOLAR, Society for Learning Analytics and Research. https://doi.org/10.18608/hla17.015 Shaffer, D. W. (2017). *Quantitative Ethnography*. Madison, WI: Cathcart Press.

Vandenberg, J., Zakaria, Z., Tsan, J., Iwanski, A., Lynch, C., Boyer, K. E., & Wiebe, E. (2021). Prompting collaborative and exploratory discourse: An epistemic network analysis study. *International Journal of Computer-Supported Collaborative Learning*, 16(3), 339–366.

https://psycnet.apa.org/doi/10.1007/s11412-021-09349-3

- Vásquez, C. (Ed.), (2022). *Research methods for digital discourse analysis.* London: Bloomsbury Publishing.
- Vujovic, M., Amarasinghe, I., & Hernández-Leo, D. (2021). Studying collaboration dynamics in physical learning spaces: Considering the temporal perspective through epistemic network analysis. *Sensors*, 21(9), 2898. https://doi.org/10.3390/s21092898

Yi-Chin Hsieh, Alvin Ping Leong

Nanyang Technological University, Singapore

yihsieh@ntu.edu.sg, alvin.leong@ntu.edu.sg

Investigating college students' perceptions of online and offline review modes in academic writing courses

Bio data



Hsieh Yi-Chin is a lecturer at the Language and Communication Centre, Nanyang Technological University, Singapore. Her current research focuses on the effectiveness of online and offline peer review practices, learning transfer in EAP courses, and English-medium instruction (EMI) in higher education. She has published in *Computer-assisted Language Learning*, *Assessment & Evaluation in Higher Education*, and *Journal of English for Academic Purposes*.



Alvin Leong is a senior lecturer at the Language and Communication Centre, Nanyang Technological University, Singapore. He taught for seven years at the National Institute of Education, and for a year at the National University of Singapore. Alvin has research interests in systemic functional grammar, discourse analysis, and literacy studies. He is the author of Theme and Rheme (Peter Lang, 2004) and a co-editor of Transforming Literacies and Language (Continuum, 2011).

Abstract

The importance of peer review practice in writing courses has been strongly supported by pedagogical research due to its value in facilitating students' writing progress. This study investigated college students' experiences with different peer review modes in an academic writing course. We investigated three peer review modes: (i) face-to-face peer review (F2F), (ii) anonymous computer-mediated peer review (CMPR), where students provided feedback anonymously on an online platform, and (iii) blended peer review, where students conducted both F2F and CMPR. This study was guided by the guestion: What are students' perceptions of and experiences with the three peer review methods? Three classes (n = 66) enrolled in an academic writing course at a Singaporean university participated in this study. The three classes were assigned to the three peer review modes respectively over a semester. Surveys and interviews were administered to investigate students' perceived usefulness of the feedback and their interactions with the reviewers. The findings show that students in all three groups were generally satisfied with the mode they were assigned to, but with a preference for the blended mode. The blended mode accommodates different learning needs by addressing the limitations of both F2F and CMPR and leveraging the merits of both modes. Several psychological and contextual factors were found to impact the effectiveness of peer review practices, including the closeness among peers, the presence of incentives, the functionality and affordances of an online peer-review platform, time constraint, and the instructions.

Proceedings of the XXIst International CALL Research Conference

Conference paper

Introduction

Peer review, where students evaluate and comment on each other's work, has been widely applied in various educational contexts due to its numerous pedagogical benefits (Cho & Cho, 2011; Cho & MacArthur, 2011, Topping, 2009). For instance, peer feedback greatly helps learners identify issues with clarity and prompts divergent thinking (Hsieh & Hill, 2021), facilitates learners' language improvement (Chang, 2014), cultivates learners' awareness of audience when they write (Lee, 2015), and engages learners in more reflective and deep thinking (Zheng et al, 2015). Apart from the focus on its benefits, another important strand of research has focused on the communication modes in the peer review process (e.g., face-to-face [F2F] and computer-mediated peer review [CMPR]). Due to the advancement of technology, a plethora of platforms and features facilitating peer review practices are made available and have become ideal alternatives to the traditional F2F peer review. Literature has shown that different peer review modes could greatly impact students' feedback-providing behaviors, the nature of feedback, students' perceptions, and the level of engagement (Chang, 2012; Ho, 2015). While most of the studies have compared the F2F and CMPR modes (Ho, 2015, Liu & Sadler, 2003; Pritchard & Morrow, 2017), very few have examined the effectiveness of a blend of the two modes and how it could impact the peer review process and student perceptions (Chang, 2012). This study investigated students' perceptions of the two primary modes of peer review, F2F and anonymous CMPR, as well as a blended mode (F2F and CMPR), in a university-level academic writing setting.

Literature review

Research has shown that both F2F and CMPR have their limitations and strengths that could impact students' perceptions and feedback-providing behaviors to a large extent.

Students' perceptions of the F2F review mode

In a F2F setting, the most valued aspect is the interactivity among students as they are given opportunities to negotiate and elaborate their intended meanings in their writing with the reviewer for informed revision decisions (Tsui & Ng, 2000). Conversely, some students in the F2F setting reported discomfort and embarrassment when they had to highlight weaknesses in each other's writing. Thus, the social factor of face-saving is found to override the provision of honest and direct comments (Bradley, 2014; Snowball & Sayigh, 2007). The fear of harming others' self-esteem and the intention to maintain cohesion may lead to poor-quality or praise-based feedback, which to some extent compromises the reliability of peer feedback. Students may also feel uncomfortable when their writings are read and critiqued by others (Topping, 2009). These feelings are summarized as "negative social processes" (Topping, 2009, p. 24), which could potentially subvert the reliability and validity of peer feedback.

Students' perceptions of the anonymous CMPR review mode

In the recent two decades, the proliferation of online peer review platforms has created alternative modes of peer review with many facilitating features, including anonymity. Technology not only streamlines the administration of anonymous peer review, but addresses the concerns in the traditional F2F peer review setting. Research has shown that anonymous CMPR has the potential to eliminate possible stress and embarrassment inherent in the F2F setting, which helps to promote more direct and honest feedback (Bradley, 2014; Loretto et al., 2016; Wu et al., 2015), as well as encourage participation due to reduced peer pressure (Raes, 2015). Additionally, when identities are unknown to each other, not only do reviewers produce more objective feedback, but writers tend to

assess the feedback more objectively (Cote, 2014). Therefore, students reported favourable attitudes toward the anonymous peer review environment (Mostert & Snowball, 2013). With the increased comfort level, Li (2017) found that students undergoing the anonymous review process academically outperformed those who engaged in fully identifiable review. However, students' chances for immediate face-to-face interactions are minimal in such an environment. Studies also found that with the nature of anonymity, some students may submit poor-quality drafts as the "peer pressure" of performing well is absent (Mostert & Snowball, 2013). The lack of engagement observed in anonymous CMPR may lead to the reviewers' reluctance to provide constructive or explicit feedback.

Research gaps and research question

These mixed results motivated us to propose a blended mode of peer review, where both F2F peer review and anonymous CMPR are implemented in the same peer review process for the same assignment. Research has been conducted to compare the F2F peer review and CMPR (Ho, 2015; Liu & Sadler, 2003; Pritchard & Morrow, 2017), yet there is a lack of studies on the blended mode. Chang (2012) is one of the few studies that examined a blended mode, with F2F as well as synchronous and asynchronous CMPR implemented in the same assignment and has concluded that different peer review modes can successfully complement each other. However, the scope of the study does not include the effect of anonymity in a blended mode. Addressing these research gaps, this study investigated students' perceptions of the effectiveness of three peer review modes: the F2F mode, anonymous CMPR, and the blended mode. Given the limitations and strengths of both the F2F peer review and anonymous CMPR mode, this study hypothesizes that a blended mode could allow the students to enjoy the merits of both- opportunities for negotiation and discussions, as well as a 'safer' environment to provide more direct and honest feedback. It is also expected that the blended mode would allow both to compensate for each other's limitations. The research question that quides this study is: What are university students' perceptions of the three peer review modes in an academic writing context?

Methodology

Study setting and participants

This study was a form of classroom action research, which is an ideal methodology that allows the teacher-researcher to test and reflect on pedagogical practices to improve teaching and learning (Creswell, 2012). The study was conducted in a compulsory academic writing course for freshmen at a Singaporean university; the course involved two-hour weekly tutorials over 12 weeks. Students from three classes participated in this study (a total of 66 students). They are all first-year students from the school of Engineering, and the male to female ratio was roughly 2:1. The three classes were randomly assigned to one of the three peer review modes: the F2F mode, the anonymous CMPR mode, and the blended mode. One of the authors was the instructor of the classes. Students were provided with an overview of the research and were fully informed of the voluntary nature of the study. Informed consent was sought from all participants.

Procedure

In this course, students wrote two assignments over a semester, a technical proposal (group work), and an evaluation report (individual work). As a usual practice in this course, students conducted a peer review with two reviewers one week prior to the submission of their final assignment draft. For the F2F group, they were given one hour to review two peers' work, write their comments, and orally discuss the feedback in class. For the CMPR group, they conducted a double-blind review in an online anonymous

platform *Peergrade*. Students were randomly assigned two drafts to review at their own time within two days. *Peergrade* was employed due to its functions that support a good practice of peer review, including anonymity, multiple reviewers, feedback reactions, etc. During the review period, the instructor monitored reviewers' feedback activities and progress. As for the blended group, students worked with one reviewer face-to-face in the class, followed by the other review conducted online over *Peergrade* (see Fig.1 for the study design). Students in all three modes conducted peer reviews for the two written assignments, as a group work for the technical proposal, and individually for the evaluation report. The same rubric was used for the assignments in all three modes. In the F2F setting, the rubric was shown on the PowerPoint slide in class, and for the CMPR, the rubric was presented along with the drafts on the screen. After the peer review, students had a week to revise their drafts based on the feedback they received from their two reviewers.



Figure 1. The study design

Data collection and analysis

To understand students' overall perceptions and learning needs, a survey was administered upon the completion of the peer review for each assignment. The reason for the repeated survey distribution is to obtain a more comprehensive observation of students' perceptions over the semester. Students were surveyed about their satisfaction in terms of the overall peer review experiences, interactions with peers, and perceived helpfulness of the feedback. To obtain a more in-depth understanding of students' perceptions, individual interview or focus-group discussions (depending on students' availability) were conducted with volunteer students from each class at the end of the semester. A total of 34 students from three of the modes participated in the interview/focus-group discussions (F2F: n=9, CMPR: n=12, blended: n= 13). In the interview, students were asked to recall the review process, peer discussions, revision decisions, and were asked about the perceived benefits, and issues with the assigned peer review mode, the overall review experiences, and their preferred review mode. The interviews were transcribed using the intelligent verbatim approach (leaving out repeated words or fillers). The interview transcripts were coded into negative and positive views using an inductive coding approach that identifies emerging common patterns and recurring themes.

Findings and discussion

The survey results and the interview/focus-group discussions are presented in a combined manner to show students' overall perceptions and preference. Exemplary quotes from the interview/focus-group discussion are provided to support the observed themes.

Overall satisfaction

The overall survey results indicate that students in all three groups were generally satisfied with the mode they were assigned to. No statistical significance was found among the three groups in terms of students' satisfaction. However, the blended mode receives the highest percentages of 'agree' and 'strongly agree' (with one exception) when students were asked about their overall satisfaction with the peer review mode for both assignment 1 and assignment 2 (Table 1).

Table 1. Percentages of 'Agree' and 'Strongly agree' in students' overall satisfaction with the peer review mode

Overall satisfaction with the peer review mode	F2F (n=23)	CMPR (n=17)	Blended (n=26)
Assignment 1	83%	82%	F2F: 88% CMPR: 88%
Assignment 2	71%	86%	F2F: 88% CMPR: 85%

The observation could be substantiated with the focus-group discussions, which revealed that the respondents generally appreciated the feedback experience. However, while the students appreciated their assigned feedback modes, they generally preferred the blended mode instead. Fifteen respondents expressly indicated their preference for the blended approach, with five other students expressing their tacit agreement through head nodding and facial gestures. Only four respondents preferred the CMPR mode and another four respondents, the F2F mode.

The respondents also raised two core issues related to their feedback experience; these pertain to the value of the feedback received, and the flexibility offered by the feedback mode. These issues are discussed in turn in the following sub-sections.

Value of feedback

Repeatedly, the respondents emphasized the importance of receiving constructive feedback. They agreed that the success of the feedback practice was viewed almost entirely in terms of how useful the feedback was in improving their drafts. Such feedback tended to be facilitated by the online mode, where the anonymity of the reviewer often resulted in feedback that was more direct and honest, which echoes the results of previous research (Bradley, 2014; Loretto et al., 2016). One student commented:

I think, being anonymous, kind of, helps with being more direct because if you know the person, you might try to like hold back your comments in case you hurt their feelings halfway. So, it's nice that it's anonymous and you don't really know who you're giving feedback to so you can be more direct about their mistakes, yah. (O1)

By contrast, the feedback received during F2F sessions was far more reserved, in accordance with many studies (Bradley, 2014; Snowball & Sayigh, 2007). As not all student-reviewer pairs were familiar with each other, the reviewers tended to be more polite and often hedged their feedback so as not to cause offence. As one respondent put it:

Yeah, it's mainly the word choices, but then I feel that at the end of the day, the word choice used may still affect the type of message that you're going to put across. Yeah, because some tone [*sic*], some words, some words are

much more straightforward and direct and hurtful. So, you will try to use less impactful words; then it may actually turn the message slightly but I will not say entirely. Alright, so message-wise, it's still a lot of the same direction, but it wouldn't be as impactful. (F15)

... when we are giving a [*sic*] feedback during face-to-face, sometimes when you want to comment on something that is, you think, your peer is doing badly and you don't dare to say. (B5)

While such a strategy may be tactful, the obvious shortcoming is that flaws in the student's draft may sometimes be glossed over or even ignored, thus affecting the value of the feedback. This problem is, of course, less severe among friends who know each other well, but such a pairing option is not always possible, particularly when greater interactivity among students in class is encouraged.

There are no easy solutions to such a problem in the F2F setting, compounded perhaps by the relative conservatism in Singaporean culture (Mathew et al., 2021). While teacher-student consultations are often held face-to-face, teachers are generally regarded as authority figures and experts in the field. But this is not so among peers. In this respect, the online mode appears to work better in helping reviewers to provide objective and potentially useful feedback.

Another issue raised by a few respondents in a F2F setting is that sometimes oral clarification may falsely indicate that an issued was solved. Some pointed out that when they clarified a problem with their peer, who may seem convinced and let the issue pass, they would still be concerned, or even confused, whether there is a need to revise their writing. One indicated, "[...] there will come this worry when others read it, will they also see this point of view or will they also need this clarification?" (B8). If they were to read the comments online, they would tend to think more objectively and thoroughly about whether, or how to revise. Some respondents even remarked that, for reviewers, such oral clarification could be a "distraction" that skewed their judgement or interrupted their concentration. These issues, again, would compromise the value of peer feedback to a great extent, and the online mode is seen as a more favorable environment for peer review. As B2 noted, "[...] in a sense doing it online removes all these distractions[...]

Time and physical constraints, and the importance of interactivity

The respondents also discussed how the time constraint on the reviewers to complete their feedback, and the need for them to be physically present during the feedback process could affect the value of the feedback. Together, their views call for a feedback mode that is flexible enough for these constraints to be relaxed.

As regards the time constraint, the F2F mode clearly suffers since the reviewer needs to read the draft and process it within the assigned time in the presence of the student. The respondents in the focus-group discussions recounted two outcomes with such a practice. First, the draft was read too quickly, resulting in feedback that was too general and thin on details. Second, conversely and more commonly, reviewers took too long to read the draft, leaving hardly enough time for oral feedback and discussion. In both outcomes, the value of the feedback is compromised. As one respondent complained:

[...] we needed to take some time to, sort of, understand what they were writing and dissect it based on the rubrics. So, we didn't really spend much time to think about, you know, whether or not the idea actually is feasible, whether it makes sense for them to continue on this idea. (F19)

The online mode, on the other hand, has no such time or physical constraints. Reviewers were allowed to read their peers' drafts and complete the feedback at their own time and convenience. This allowed them to think more deeply about both the content and technical aspects of their assigned drafts, and so offer more detailed, constructive, and valid suggestions for improvement. Coupled with the functionalities and affordances of the online review platform, students' review practices were further facilitated. Having the paper and the rubric juxtaposing on the screen "prompts us to check through the points," O2 indicated, therefore encouraging more comprehensive feedback. Reviewers also felt incentivized by the feedback generated by the system for the reviewers. "[...] like 'Wow, you're a feedback rock star' [...] that's quite helpful and it provides people some motivation to give more elaborate feedback. (O3)" These functionalities helped to renduce the positive experiences in a constraint-free online review setting and are not readily available in a F2F setting.

While the time and physical constraints of the F2F mode might appear to be hindrances, numerous respondents in fact valued the physical presence of the reviewer. When it worked well, the F2F mode allowed for immediate interaction between students and reviewers; the ease and convenience of speaking, rather than typing, about points of concern was especially attractive. F22, in fact, described his feedback experience as "a very intimate feel", and appreciated how the interactivity allowed him to absorb the information more quickly. Seeking such immediate clarification in the online mode is not quite possible, and while there is an option to deliver an online message to the reviewer via the feedback reactions function, the message is not always answered promptly.

[...] unlike the comment part where it is required for us to comment on the friend's work on Peergrade. But to reply is not a must. So, we will choose the easy option, which is not to reply. (B22)

Not replying to a query defeats the idea of feedback being "a dialogical and contingent two-way process" (Nicol, 2010, p. 503). This is an important point to consider, for while the feedback offered via the online mode tends to be more detailed, it can also be challenging. In the event that clarification is needed, the online mode lacks the synchronous interactivity that the F2F mode offers.

Proposed blended peer feedback model

The inputs of the respondents indicate clearly that feedback practice, including the mode of feedback, is only as useful as the perceived value of the feedback received. The respondents preferred peer feedback to be honest and constructive, and expected the peer review process to be free from "distractions." They also highlighted the need for the feedback mode to be flexible enough to facilitate both objective feedback and synchronous interactivity between students and reviewers. The following model summarizes the broad observations from the interview and focus-group discussions:



Figure 2: Proposed blended peer feedback model

We argue that the peer feedback mode that is best able to achieve this flexibility is the blended mode, where both the F2F and the CMPR modes are incorporated. The blended approach not only brings together the strengths of each individual mode to accommodate the different learning needs of individual students, but also addresses the limitations of both modes. Some students may have a greater preference for one or the other mode, and may thus fail to fully benefit from the feedback process if only the non-preferred mode is used. This problem is mitigated if both modes are used.

Further, for any single assignment, the use of both modes has the added advantage of including more reviewers — one for the F2F component, and another for the online component. This helps students to consider the perspectives and recommendations from more than one reviewer, and so make better decisions to revise and refine the writing (Cho & Schunn, 2007). In fact, four respondents, such as F22 below, recommended having more than two reviewers despite knowing that this would entail more work for themselves.

Researcher:	So, when you say "more of that", do you mean more
	reviewers or multiple times of review?
F22:	More reviewers. [] I think two is the bare minimum.
Researcher:	Oh, I see. So, maybe three or even four would be better?
F22:	Yeah, three or four will be better. [] When you go out to work, you have to work with three or four people so it will be better to work with that number of people so you can train your mind. (F22)

We see here an implicit benefit of the feedback process, that it not only helps students improve on their writing, but goes some way to prepare them for the future demands of the workplace, where working in teams and providing feedback are essential skills (Krakoff, n.d.). Particularly, to be able to provide feedback via different modes and platforms, the provision of modelling in the instruction is important to prepare the students for feedback-providing and to maximize the efficacy of a blended peer review mode.

Conclusion

This study investigated university students' perceptions of different modes of peer review, F2F and anonymous CMPR, and a blended mode. The findings largely support our

Proceedings of the XXIst International CALL Research Conference

hypothesis that the blended mode is perceived to be the most conducive practice that allows students to enjoy the merits of both modes. We proposed a peer feedback model, in which both F2F and anonymous CMPR are incorporated to create the desired flexibility that facilitates peer review practices.

References

- Bradley, L. (2014). Peer-reviewing in an intercultural wiki environment-student interaction and reflections. *Computers and Composition, 34*, 80-95. https://doi.org/10.1016/j.compcom.2014.09.008
- Chang, C. (2012). Peer review via three modes in an EFL writing course. *Computers and Composition*, 29(1), 63–78. https://doi.org/10.1016/j.compcom.2012.01.001
- Chang, C. (2014). *L2 peer review: Insights from EFL Taiwanese composition classrooms*. Crane.
- Cho, Y. H., & Cho, L. (2011). Peer reviewers learn from giving comments. *Instructional Science*, *39*(5), 629–643. https://doi.org/10.1007/s11251-010-9146-1
- Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology*, 103(1), 73-84. https://doi.org/10.1037/a0021950
- Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the disciplines: A web-based reciprocal peer review system. *Computers & Education*, 48(3), 409–426. https://doi.org/10.1016/j.compedu.2005.02.004
- Cote, R. A. (2014). Peer feedback in anonymous peer review in an EFL writing class in Spain. *Gist Education and Learning Research Journal*, 9, 67-87. http://files.eric.ed.gov/fulltext/EJ1062639.pdf
- Creswell, J. (2012). Educational research. Pearson Education.
- Ho, M. C. (2015). The effects of face-to-face and computer-mediated peer review on EFL writers' comments and revisions, *Australian Journal of Educational Technology*, 31(1), 1-15. https://doi.org/10.14742/ajet.495
- Hsieh, Y. C. & Hill, C. (2021). Reconceptualizing the value of peer and instructor feedback using a sequential structure, *Assessment & Evaluation in Higher Education*. Advance online publication. https://doi.org/10.1080/02602938.2021.1998341
- Krakoff, S. (n.d.) *How to give constructive feedback in the workplace*. Champlain College Online. Retrieved June 2, 2022, from
- https://online.champlain.edu/blog/giving-constructive-feedback Lee, M. K. (2015). Peer feedback in second language writing: Investigating junior secondary students' perspectives of inter-feedback and intra-feedback. *System*, 55, 1-10. https://doi.org/10.1016/j.system.2015.08.003
- Li, L. (2017). The role of anonymity in peer assessment. Assessment & Evaluation in Higher Education, 42(4), 645-656.
 - https://doi.org/10.1080/02602938.2016.1174766
- Liu, J., & Sadler, R. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes, 2*(3), 193–227. https://doi.org/10.1016/S1475-1585(03)00025-0
- Loretto, A., DeMartino, S., & Godley, A. (2016). Secondary students' perception of peer review of writing. *Research in the Teaching of English*, 51(2), 134–161. https://www.jstor.org/stable/24889912
- Mathew, M., Teo, K. K., Tay, M., & Wang, A. (2021). Our Singaporean values: Key findings from the World Values Survey. Institute of Policy Studies. https://lkvspp.nus.edu.sg/docs/default-source/ips/ips-exchange-series-16.pdf
- Mostert, M., & Snowball, J. D. (2013). Where angels fear to tread: Online peer assessment in a large first-year class. *Assessment & Evaluation in Higher Education, 38* (6), 674-686. https://doi.org/10.1080/02602938.2012.683770
- Nicol, D. (2010). From monologue to dialogue: Improving written feedback processes in mass higher education. Assessment & Evaluation in Higher Education, 35(5), 501–17. https://doi.org/10.1080/02602931003786559

- Pritchard, R. J., & Morrow, D. (2017). Comparison of online and face-to-face peer review of writing. *Computers and Composition*, 46, 87–103. https://doi.org/10.1016/j.compcom.2017.09.006
- Raes, A. (2015). Increasing anonymity in peer assessment by using classroom response technology within face-to-face higher education. *Studies in Higher Education*, 40(1), 178–193. https://doi.org/10.1080/03075079.2013.823930
- Snowball, J. D., & Sayigh, E. (2007). Using the tutorial system to improve the quality of feedback to students in large class teaching. South African Journal of Higher Education, 21(2), 321–333. https://hdl.handle.net/10520/EJC37323
- Topping, K. J. (2009). Peer assessment. *Theory into Practice, 48*(1), 20-27. https://doi.org/10.1080/00405840802577569
- Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? Journal of Second Language Writing, 9, 147-170. https://doi.org/10.1016/S1060-3743(00)00022-9
- Wu, W. C., Petit, E., & Chen, C. H. (2015). EFL writing revision with blind expert and peer review using a CMC open forum, *Computer Assisted Language Learning*, 28(1), 58-80, https://doi.org/10.1080/09588221.2014.937442
- Zheng, B., Lawrence, J., Warschauer, M., & Lin, C. H. (2015). Middle school students' writing and feedback in a cloud-based classroom environment. *Technology*, *Knowledge and Learning*, 20, 201–229. https://doi.org/10.1007/s10758-014-9239-z

Yin Yang, Yanjie Song

The Education University of Hong Kong, HKSAR, China

yyin@s.eduhk.hk, ysong@eduhk.hk

A study on social interactions among primary students in English vocabulary acquisition in a mobile learner-generated content learning environment

Bio data



Yin Yang is a PhD candidate at The Education University of Hong Kong. Her research interests include technology-enhanced language learning and second language acquisition. She received the Best Teacher Forum Paper Award in GCCCE 2019 and Best Oral Presentation Award in PPTELL 2021.



Yanjie Song is an associate professor in the Department of Mathematics and Information Technology at The Education University of Hong Kong. Her research areas focus on mobile learning, seamless learning, technology-enhanced language learning (e.g. augmented reality), vocabulary acquisition, teacher orchestration with technology, learning analytics, innovative pedagogy design, and implementation.

Abstract

This article reports on a study exploring primary students' after-class social interactions in English vocabulary acquisition using a mobile learner-generated-content (m-LGC) tool. A total number of 29 grade 4 students from an elementary school in Hong Kong were involved. A case study approach was adopted. Data collection included log data on the m-LGC tool and semi-structured interviews. Data analysis included content analysis, visualisation using Gephi, and thematic analysis. The results show two types of students' social interactions in a m-LGC learning environment. The interview results indicate that students held different perceptions of social interactions using the m-LGC tool. The implications are discussed.

Conference paper

Introduction

The development of mobile technologies provides learners with a collaborative, flexible, real-time learning experience in which learners can study a second language regardless of time and location constraints (Jeong, 2022; Sung et al., 2015). Many studies suggest that mobile-assisted language learning (MALL) facilitates interactions and deeper integration of language learning with real-life needs and cultural experiences (Chen, 2016; Derakhshan & Khatir, 2015). In Long's Interaction Hypothesis (1996), interactive

tasks promote negotiation of meaning among learners, which can facilitate the development of a second language via connecting input, internal learner capacities and output in productive ways. However, few studies have been conducted which explore social interactions supported by mobile devices in English vocabulary acquisition.

This study adopted a case study approach to explore social interactions among Hong Kong primary students in English vocabulary acquisition within a mobile learner-generated content (m-LGC) learning environment.

Literature review

The sociocultural theory was proposed by Vygotsky (1978) that stressed the role of community in the process of "making meaning". Learning was at first social, later individual. The zone of proximal development (ZPD) occurs when the social interaction occurs between a student and a more knowledgeable individual (Vygotsky, 1978). Social networking tools provide extraordinary opportunities for students to make connections with peers and teachers regardless of time and location constrains (Akbari et al., 2015; Mellati et al., 2018).

However, previous studies have mainly addressed face-to-face interaction in second language acquisition (Barnes et al., 2017; Tratnik et al., 2019). Some have investigated the effects of peer interaction supported by digital technologies on second language vocabulary learning (Verga & Kotz, 2017; Mellati et al., 2018), but most of them were conducted using prescribed learning tasks (e.g. a vocabulary learning game designed by researchers, see for example Verga & Kotz, 2017), or only in classrooms (e.g. Mellati et al., 2018). Few studies have gained an understanding of how learners use digital tools beyond the classroom (e.g. Lai & Zheng, 2018) and how social interactions happen in user-generated learning content mediated by digital technology. In addition, only a few studies integrated social interaction with the curriculum to support learning beyond the classroom (Richards, 2015).

Research aims and questions

This study aimed to understand how primary students use the mobile learner-generated content (m-LGC) tool in peer-to-peer interactions, while engaging in a user-generated learning environment outside of the language classroom. The following research questions were addressed:

- RQ1. What types of social interactions among students occurred in a mobile-user-generated-content (m-LGC) learning environment?
- RQ2. What was the student perception of involving social interactions during vocabulary acquisition in the m-LGC learning environment?

Research design

The tool used in this study

The mobile learner-generated content (m-LGC) tool used in this study was adapted from SCROLL (System for Capturing and Reminding of Learning Log) (see Ogata et al., 2011; Song & Ma, 2021; Song & Yang, 2019). Figure 1 shows the interface of the m-LGC tool on mobile devices. Learners can create a learning log via taking or uploading pictures, inputting the target word, describing the newly acquired vocabulary and making an audio recording. Learners can also input synonyms and collocations of the target words.



Figure 1. The interface of creating a "learning log"

The function used in this study focused on social interactions supported by a commenting function in Learning Community (refer to Figure 2). The m-LGC tool provides students with opportunities to learn from peers by giving text-based comments. Learners can not only generate their learning logs but also reflect on how peers describe the words and use the tool to discuss new words they encounter in real life. In Learning Community, learners can make comments on peers' learning logs.



Figure 2. Social interactions among students in Learning Community on the m-LGC tool

Proceedings of the XXIst International CALL Research Conference

Participants

A case study approach was adopted in order to uncover students' social interactions in English vocabulary acquisition in an m-LGC learning environment (Yin, 2002). Twenty-nine Grade 4 students (14 females and 12 males) aged between 10 and 11 in a primary school from Hong Kong were involved in this study. To consider the research ethics of a study that involved collecting data from the participants, a written informed consent form was obtained from both the participants and their parents.

Instructional design

The study lasted for two weeks. Before performing the learning task, the teacher briefed the students on how to use the m-LGC tool. Students could practice and consolidate the words at home. The topic of vocabulary learning reported in this paper was 'a healthy life'.

Data collection and analysis

Data collection included (1) log data on the m-LGC tool: raw event log data stored in MySQL database (for importing spreadsheet with relationship information, e.g. source, target), and students' comments on learning logs; and (2) semi-structured interviews: twelve students were invited to understand their perceptions of involving social interactions during vocabulary acquisition in the m-LGC learning environment.

Data analysis included content analysis, visualisation of network graph using Gephi and thematic analysis. To address RQ1, content analysis was adopted to analyse the students' comments. A coding scheme modified from Shea et al. (2010) was adopted to analyse types of students' comments on the m-LGC tool in terms of two dimensions (see Table 1): affect (AF) and open communication (OC). Two coders were involved in analysing the types of students' comments. The inter-rater reliability for the results of coding was 0.93. All the discrepancies were discussed and solved. Then, Gephi (https://gephi.org), an interactive visualisation platform, was used to visualise students' social connections on the m-LGC tool. The visualisation was done using Gephi's layout algorithm (ForceAtlas2), which was suitable for small sample sizes (Jacomy et al., 2014; Khokhar, 2015). The features of two types of students' social interactions were discussed. To address RO2, thematic analysis was used to analyse focus group interviews. The interviewees were selected based on the frequency of students' comments on peers' logs. Students with high and low frequencies in each AF and OC social interactions were identified and invited. In total, twelve students were involved in the focus group interview. The interview lasted for 45 minutes and was recorded for further analysis. The interview was conducted in Cantonese and was transcribed into English.

Results

Types of social interactions in a m-LGC learning environment

To analyse the types of social interactions in an m-LGC learning environment, firstly, students' comments logged on the m-LGC tool were retrieved and categorised using content analysis. Secondly, students' log data on the m-LGC tool was transformed into data that could be read by the visualisation tool Gephi to show the overall picture of social interactions of students. The results are presented below.

Categories	Items	Definition	Examples
Affect (AF)	expression of feelings	expressions of emotion, includes repetitious punctuation and emoji	e.g., I am excited to see this picture.
	self-disclosure	present details of life beyond the class; includes expression of likes or dislikes	e.g., I love eating sandwiches.
	expression of values	express personal values and beliefs	e.g., I feel our children should not eat junk food.
Open communication (OC)	asking questions	ask questions of other students	e.g., Do you know how to make sushi?
	answering questions	answer questions of other students	e.g., You can go to the websiteto find more information:)
	referring explicitly to target learning logs	further illustrations of the learning logs using the target words	e.g., Curry beef is yummy. (<i>Note: Curry</i> <i>beef is the</i> <i>target word</i>)
	Expressing agreement/disagreement	Expressing agreement/disagreement with others or the descriptions of learning logs	e.g., You are right/I don't think

Table 1. Coding scheme of types of social interaction

A total of 93 students' comments were collected in this study, of which 66 comments were classified as AF and 27 as OC. Figure 3 shows the distribution of students' social interactions in the category of AF. In the category of AF, 48.5% of comments (n=32) were related to expressing feelings using emojis and words expressing likes or dislikes (e.g. Happy!/Good!/ ($\ge \omega \le$)/), 31.8% (n=21) presenting details of life beyond the class (e.g. I love eating vegetables), and 19.7% (n=13) expressing personal values and beliefs (e.g. You can't eat too much junk food; Dairy products is good for your health.)





Figure 4 shows the distribution of students' social interactions in the category of OC. In

the category of OC, 48.1% of the comments (n=13) were related to further illustrating the learning logs (e.g. I see tomatoes/ I put an apple in the fridge), followed by 22.2% of comments (n=6) asking questions (e.g. Do you like drinking it?/ you make??) and 18.5% of comments (n=5) expressing disagreements or agreements (e.g. Me too/Yes). Only three comments were answering questions raised by other students (e.g. Yes, I like drinking apple juice).



Figure 4. Distribution of students' social interactions in the category of OC

Figures 5 and Figure 6 show students' social interactions in a m-LGC learning environment in terms of categories AF and OC, respectively. The size of the node represents the 'degree centrality' (its number of connections). The larger the node, the higher degree of centrality it represents. The thickness of the edge represents the frequency of students' comments on peers' logs. The thicker the edge, the higher frequency it represents. An arrow indicates the direction of the target. For example, as showed in Figure 5, the arrow between the node of 'Anna' and the node of 'Bob' indicate that Anna made comments on Bob's learning logs. The number on the node represents in total.

Figure 5 shows that in the category of AF, students were engaged in making comments on peers' learning logs in general. According to the size of the node, Jim, Penny, Tim, Floria, and Charles had a higher number of received comments than other students. According to the thickness of the edge, Anna, Adora, Jim, Mendy, Sharon, Tim and Charles had a thinker edge than others, indicating that these students were active in making comments.



Figure 5. Visualisation of students' social interactions in the category of AF

Figure 6 shows the visualisation of students' social interactions in the category of OC. Compared with students' social interactions in the category of AF, students' social interactions in the category of OC were less frequent; only 14 students were involved in posting comments.

According to the size of the node, Adora, Tim, and Anna had a higher number of received comments than other students. According to the thickness of the edge, Anna, Adora, Tim had a thinker edge than others, indicating that these students were active in making comments.



Figure 6. Visualisation of students' social interactions in the category of OC

Proceedings of the XXIst International CALL Research Conference

Students' perceptions of social interactions during vocabulary learning in the m-LGC learning environment

The focus group interviews indicate that students held different perceptions of social interactions during vocabulary learning in the m-LGC learning environment. The thematic analyses identified factors that influenced students' engagement in posting comments in the m-LGC learning environments as follows.

High frequency of social interactions

For student who were active in posting comments to express their feelings, likes or dislikes and values and beliefs, or further illustrate the learning logs, or ask/answer questions, the three most mentioned reasons were identified:

- Willingness to share personal feelings: Many students said they were willing to share personal feelings with others. For example, one interviewee explained: "When I found interesting logs posted by my classmates, I would make comments. I would like to share my feelings with others." (Anna)
- **Positive mindset:** Students mentioned that they did not care about what other classmates thought of their comments. They tended to make positive comments, hoping to encourage classmates. For example, one interviewee explained: "I hope my positive comments can make my classmates feel happy and motivated." (Charles)
- **Treating making comments as a learning opportunity:** Students stated that they would make comments by using the target words to memorise them better. For example, two interviewees explained:

"I love viewing peers' learning logs. I asked myself about how I would describe the picture posted. So, I made sentences using the target words again in comments." (Andora)

"I just felt the learning experience was interesting and novel." (Tim)

Low frequency of social interactions

For students who were not active in posting comments to express their feelings, likes or dislikes and values and beliefs, or further illustrate the learning logs, or ask/answer questions, the three most mentioned reasons were identified:

• Fear of losing face: Students mentioned they were afraid of losing face. 'Face' is typical cultures in East Asian countries. People try to avoid being embarrassed. For example, one interviewee explained:

``It would be very embarrassed if I made inappropriate comments. Classmates may laugh at me." (Martin)

- ٠
 - **Fear of hurting classmates:** Students mentioned they cared about what other people may think of their comments. They tried not to make others feel uncomfortable. For example, one interviewee explained:

"I was afraid that my true feelings may hurt my classmates. For example, some sentences were full of spelling mistakes. If I pointed them out, my classmates may feel embarrassed. So, I did not make any comment." (Peter)

• Less motivation: Students mentioned they were bored of making comments because there were no rewards or benefits. For example, one interviewee explained:

"Making comments was a waste of time. There was no reward; why should I do that?" (Sandy)

Conclusion and implications

The results of the study identified two types of students' social interactions in an m-LGC learning environment. Students were more engaged in posting comments related to the category of AF than posting comments related to the category of OC. In addition, the overall picture of students' social interactions in two types were visualised using Gephi. Students with high and low frequencies in each AF and OC social interactions were identified. The interview results indicate that students held different perceptions of social interactions during vocabulary learning in the m-LGC learning environment.

Considering the tentative nature of this study, this study has its limitations. Frist, this study had a small sample size and a short study duration. Thus, future research is needed to include a larger sample size over a longer duration. Second, the coding in this study was labour-intensive. In the future, text-mining techniques could be used to identify the characteristics of students' social interactions. Third, this study did not examine the relationship between students' social interaction patterns with their vocabulary learning performance. Future studies could be conducted in that direction.

The significance of this study lied mainly in three aspects: (1) it was a novel study that addressed social interactions beyond the classroom among primary students in vocabulary acquisition mediated by the m-LGC tool in Hong Kong; (2) it identified the features and patterns between social interactions and learners' vocabulary acquisition which has rarely been investigated in previous studies; and (3) the results can inform the pedagogical design of vocabulary acquisition involving students' online social interactions in a learner-generated-content learning environment. The factors influencing students' overcome negative mindset in learning.

Acknowledgements

The study was funded by the General Research Fund (Ref. 18611019), Research Grants Council, University Grant Committee, Hong Kong.

References

- Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, 3(3), 21-34. DOI: 10.52950/TE.2015.3.3.002
- Barnes, E. M., Dickinson, D. K., & Grifenhagen, J. F. (2017). The role of teachers' comments during book reading in children's vocabulary growth. *The Journal of Educational Research*, *110*(5), 515-527. https://doi.org/10.1080/00220671.2015.1134422
- Chen, X. (2016). Evaluating language-learning mobile apps for second-language learners. Journal of Educational Technology Development and Exchange (JETDE), 9(2), 3. DOI: 10.18785/jetde.0902.03
- Derakhshan, A., & Khatir, E. D. (2015). The effects of using games on English vocabulary learning. *Journal of Applied Linguistics and Language Research*, 2(3), 39-47.
- Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PloS one*, *9*(6). https://doi.org/10.1371/journal.pone.0098679

- Jeong, K. O. (2022). Facilitating sustainable self-directed learning experience with the use of mobile-assisted language learning. *Sustainability*, *14*(5), 2894. https://doi.org/10.3390/su14052894
- Khokhar, D. (2015). Gephi cookbook. Packt Publishing Ltd.
- Lai, C., & Zheng, D. (2018). Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL*, 30(3), 299-318. DOI:10.1017/S0958344017000258
- Liao, Y. W., Huang, Y. M., Chen, H. C., & Huang, S. H. (2015). Exploring the antecedents of collaborative learning performance over social networking sites in a ubiquitous learning context. *Computers in Human Behavior*, 43, 313-323. https://doi.org/10.1016/j.chb.2014.10.028
- Mellati, M., Khademi, M., & Abolhassani, M. (2018). Creative interaction in social networks: Multi-synchronous language learning environments. *Education and Information Technologies*, 23(5), 2053-2071. DOI: 10.1007/s10639-018-9703-9
- Ogata, H., Li, M., Hou, B., Uosaki, N., El-Bishouty, M., & Yano, Y. (2011). SCROLL: Supporting to share and reuse ubiquitous learning log in the context of language learning. *Research & Practice in Technology Enhanced Learning*, 6(2), 69–82. Retrieved form https://www.academia.edu/2743228/SCROLL_Supporting_to_Share_and_Re_

https://www.academia.edu/2743228/SCROLL_Supporting_to_Share_and_Re use_Ubiquitous_Learning_Log_in_the_Context_of_Language_Learning

- Richards, J. C. (2015). The changing face of language learning: Learning beyond the classroom. *RELC Journal*, 46(1), 5-22. https://doi.org/10.1177/0033688214561621
- Sung, Y.-T., Chang, K.-E., & Yang, J.-M. (2015). How effective are mobile devices for language learning? A meta-analysis. *Educational Research Review*, 16, 68-84. https://doi.org/10.1016/j.edurev.2015.09.001
- Shea, P., Hayes, S., Vickers, J., Gozza-Cohen, M., Uzuner, S., Mehta, R., & Rangan, P. (2010). A re-examination of the community of inquiry framework: Social network and content analysis. *The Internet and Higher Education*, 13(1-2), 10-21. https://doi.org/10.1016/j.iheduc.2009.11.002
- Song, Y., & Ma, Q. (2021). Affordances of a mobile learner-generated tool for pupils' English as a second language vocabulary learning: An ecological perspective. British Journal of Educational Technology, 52(2), 858-878. https://doi.org/10.1111/bjet.13037
- Song,Y., & Yang, Y. (2019, May). Bridging in-class and real life vocabulary learning using a mobile learner-generated content tool: A case study. The 23rd Global Chinese Conference on Computers in Education (GCCCE). Wuhan, China. Retrieved from https://www.academia.edu/39294558/Bridging_in-class_and_real_life_vocabular y_learning_using_a_mobile_learner-generated_content_tool_A_case_study
- Tratnik, A., Urh, M., & Jereb, E. (2019). Student satisfaction with an online and a face-to-face business English course in a higher education context. *Innovations* in Education and Teaching International, 56(1), 36-45. https://doi.org/10.1080/14703297.2017.1374875
- Verga, L., & Kotz, S. A. (2017). Help me if I can't: Social interaction effects in adult contextual word learning. *Cognition*, 168, 76-90. https://doi.org/10.1016/j.cognition.2017.06.018
- Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard University Press.
- Yin, R.K. (2002) Applications of case study research. Sage.

Yin Yang, Yanjie Song

The Education University of Hong Kong, HKSAR, China

yyin@s.eduhk.hk, ysong@eduhk.hk

Examining primary student self-regulated vocabulary learning behavioural patterns and vocabulary learning outcomes leveraged by the mobile app with a self-regulation scheme

Bio data



Yin Yang is a PhD candidate at The Education University of Hong Kong. Her research interests include technology-enhanced language learning and second language acquisition. She received the Best Teacher Forum Paper Award in GCCCE 2019 and Best Oral Presentation Award in PPTELL 2021.



Yanjie Song is an associate professor in the Department of Mathematics and Information Technology at The Education University of Hong Kong. Her research areas focus on mobile learning, seamless learning, technology-enhanced language learning (e.g., augmented reality), vocabulary acquisition, teacher orchestration with technology, learning analytics, innovative pedagogy design, and implementation.

Abstract

This research reports a case study which provides insights into primary students' SRVL processes and behaviours that interact over time in authentic learning environments beyond the classroom, leveraged by a mobile app with a self-regulation scheme. The participants were 44 grade four students in a government-funded primary school in Mainland China. Data collection included log data on the app, and pre- and post-vocabulary tests. Data analysis included clustering, progress-mining techniques, and Kruskal Wallis tests. The findings showed that (1) students' SRVL behaviours leveraged by a mobile app with a self-regulation scheme could be clustered into three groups, (2) the characteristics of SRVL behavioural patterns among three clusters were discussed, and (3) a significant association between the identified three clusters and the students' vocabulary learning outcomes was observed.

Conference paper

Introduction

It is widely acknowledged that self-regulated learning (SRL) is one of the most essential capabilities for lifelong learning to cope with the challenges of the twenty-first century (Lehmann et al., 2014; Zheng et al., 2018). A number of studies have suggested that students who can regulate language learning tend to perform better than those who do

Proceedings of the XXIst International CALL Research Conference

not engage in self-regulation (Saks & Leijen, 2019; Zheng et al., 2018). Recently, due to the advancement of mobile technologies, the vocabulary learning experience has become ubiquitous and contextualised by integrating multimedia. But only limited mobile technologies adopted in current studies appear to support the whole process of self-regulated vocabulary learning (SRVL) (Yang et al., in press). Furthermore, little is known about how mobile technologies support primary students' SRVL in forethought, performance and reflection processes (Zimmerman, 2002), and whether student SRVL behaviours are related to their vocabulary learning outcomes or not.

This study adopted a case study to get insights into the students' SRVL processes and behaviours in authentic learning environments beyond the classroom leveraged by a mobile app with a self-regulation scheme. The following research questions were addressed:

(1) What were primary students' SRVL behavioural patterns leveraged by the mobile app with a self-regulation scheme?

(2) If there were different SRVL behavioural patterns, do students with different patterns differ in their vocabulary learning outcomes?

Literature Review

Self-regulation theory

SRL refers to students' self-initiated actions involving setting goals, monitoring their efforts to achieve goals, regulating their cognitive and metacognitive processes and learning behaviours in their learning processes, and reflecting (Pintrich, 2000; Zimmerman, 2002). Zimmerman (2002) stated that SRL involved three cyclic phases, namely, forethought (e.g., goal setting, strategic planning), performance (e.g., self-observation, self-control), and self-reflection (e.g., self-judgement, self-evaluation). In the forethought phase, the students analyse the learning task which involves goal setting and strategic planning. In the performance phase, the students perform the task while monitoring the learning process. At the same time, they use self-control strategies to keep themselves engaged in learning tasks. In self-reflection, students assess their learning performance and satisfaction, evaluate the strategies used, and reflect on what they will do in the next round of learning.

Self-regulated vocabulary learning using learning analytics

To date, a growing number of studies aim at enhancing students' SRVL in real life learning settings supported by technologies. Aligned with the modern SRL research, SRL was considered as a dynamic process (Li, et al., 2020; Panadero et al., 2016). Despite this, few studies have attempted to investigate SRVL processes and behaviours utilising learning analytics. Even fewer studies have been conducted to examine the relationship between students' SRVL behavioural patterns and vocabulary learning outcomes. In most cases, researchers have taken a variable-centered approach, for example, exploring the features of SRL behaviours between learners with high and low academic performance (Yang et al., 2018). Yet only a limited number of studies have sought to examine how specific SRL behaviours cluster among individual learners (Jang et al., 2017; Li et al., 2020). Thus, there is a significant need to cluster students' SRVL behaviours by putting the students into homogenous groups with similar profiles in order to obtain insights into the generalised patterns of students' SRVL behaviours in mobile learning environments.

This study firstly identified student SRVL behavioural patterns via clustering homogenous groups with similar SRVL behaviours; then, features of each distinct SRVL behavioural pattern were explored using learning analytics. Finally, the relationship between the identified groups of SRVL behavioural patterns and vocabulary learning outcomes was investigated.

Methods

Participants

206

Participants were 44 grade 4 students in a government-funded primary school in Mainland China. The study lasted for four weeks. The learning unit reported in this study was "Dinner's ready."

Data collection and analysis

Data collection included log data on the mobile app, and pre-and post-vocabulary tests. To understand students' SRVL behaviours and processes leveraged by a mobile app with a self-regulation scheme, a theoretical lens of micro-level SRVL based on the SRL model (Saint et al., 2020; Zimmerman, 2002) was adopted. Table 1 represents students' SRVL behaviours and processes leveraged by a mobile app with a self-regulation scheme in terms of SRL cyclic phases, micro-level SRVL processes and specific activities on the mobile app. In general, SRL was composed of three cyclic phases: forethought, performance, and self-reflection. During the forethought phase, the students analysed the learning task, including goal setting and strategic planning. The app adopted in this study enabled students to set goals regarding words, learning time, and expected ranking. In addition, students could plan specific activities and adjust their learning goals. During the performance phase, students could create learning logs by taking pictures, recording and inputting words or sentences, while simultaneously monitoring their progress. Last, the students performed self-evaluation by taking quizzes, assessing their learning performance, evaluating strategies utilised, and reflecting on what they would do in the next learning cycle. To measure students' learning performance, vocabulary learning tests were used. Vocabulary tests consisted of the Vocabulary Levels Test (VLT) and self-constructed curriculum-based vocabulary learning outcomes tests. The tests were designed to examine both breadth and depth of word knowledge (Schmitt, 1999). The internal consistency has been confirmed in a pilot study with a Cronbach's alpha value above 0.80.

SRL cyclic	Micro-level	
phases	SRVL	SRVL Descriptions
Performance	Goal-setting	Students set learning goals (e.g., number of learning logs, time, ranking)
	Strategic planning	Students plan the learning strategies to reach the goal.
		Students reset learning goals.
Performance	Self-observation	Students check the overview of recorded learning logs, time spent on the app, and ranking.
	Self-control	Students post learning logs.
		Students review/edit their own logs
		Students review peers' logs
Self-reflection	Self-judgement	Students take quizzes.
		Students evaluate the performance and the efforts.
	Self-reaction	Students evaluate self-satisfaction.
		Students evaluate planned strategies.
		Students reset/modify plans.

Table 1. Overview of students' SRVL behaviours on the app

Data analysis included agglomerative hierarchical clustering, progress-mining techniques using the R package - PMineR, and Kruskal Wallis tests.

Proceedings of the XXIst International CALL Research Conference

Results

Primary students' SRVL behavioural patterns

After applying the agglomerative hierarchical clustering and using the silhouette method (Dinh et al., 2019) to choose the optimal number of clusters, three clusters were identified. Figure 1 shows the cluster plot, which groups similar SRVL behaviours using a Euclidean distance metric.



Figure 1. Cluster plot of the clustering result of primary students' SRVL behaviours

Then, the process mining technique "First Order Markov Models (FOMMs)" in the pMineR package was employed for primary students' SRVL behaviours. It showed the likelihood of transition among each micro-level SRVL process (Matcha et al., 2019; Peeters et al., 2020). The lines between one node and the next represent the transition probability (TP), which refers to a stochastic measure of the likelihood of transition between one node to another (Saint et al., 2020). The FOMM graphs were generated to compare the differences among the identified three clusters. The similarities and differences 3, and Figure 4.

As for the similarities, three aspects were identified. First, three clusters shared similar learning patterns in goal-setting, strategic planning and self-observation. For example, Figure 2 shows that students in Cluster 1 tended to initiate SRVL by setting goals with a transition probability of 0.84, making strategic planning with a transition probability of 0.25, and monitoring the learning process with a transition probability of 0.26. Figure 3 shows that students in Cluster 2 initiated SRVL by setting goals with a transition probability of 0.74, making strategic planning with a transition probability of 0.25, and monitoring the learning process with a transition probability of 0.21. Figure 3 shows that students in Cluster 2 initiated SRVL by setting goals with a transition probability of 0.25, and monitoring the learning process with a transition probability of 0.21. Figure 3 shows that students in Cluster 3 also exhibited a similar pattern, starting with setting goals with a transition probability of 0.24, and monitoring the learning process with a transition probability of 0.27. Second, students in the three clusters showed non-linear learning trajectories between goal-setting and strategic planning, indicating students would revise learning goals. Third, self-reaction was the last activity among students in the three clusters, with a transition probability of 0.41, 0.39, and 0.42, respectively.

Proceedings of the XXIst International CALL Research Conference
As for the differences, two main aspects were identified. First, compared to students in Cluster 1 starting from goal-setting, some students in Cluster 2 moved to self-control with a transition probability of 0.16 to post learning logs, and students in Cluster 3 would engage themselves in self-observation with a transition probability of 0.33. Second, the students in Cluster 2 adopted comprehensive SRL strategies with non-linear learning trajectories. Figure 3 shows that starting from the performance (self-observation and self-control) and the reflection phase (self-judgement and self-reaction), students in cluster 2 adopted comprehensive SRL strategies with various learning trajectories that were mostly non-linear and across different micro-level SRVL processes. While Figure 2 and Figure 4 show that students in Cluster 1 and Cluster 3 followed a linear learning trajectory in the performance phase (self-observation and self-control), and focused much on posting and/or viewing learning logs. Students seldom reviewed the learning process or ranking after posting/viewing learning logs.



Figure 2. First Order Markov Models (FOMMs) of Cluster 1



Figure 3. First Order Markov Models (FOMMs) of Cluster 2



Figure 4. First Order Markov Models (FOMMs) of Cluster 3

Relationship between identified clusters and vocabulary learning outcomes

A non-parametric test named Kruskal Wallis Test was adopted to examine the difference of pre-vocabulary tests among students in the three identified clusters, as the pre-test was not normally distributed. The results indicated that there was no significant

Proceedings of the XXIst International CALL Research Conference

209

difference in students' prior English levels among the three identified clusters before the study.

The scores of post-vocabulary learning outcomes for all clusters are presented in Table 2. Kruskal Wallis tests showed a significant association between the identified three clusters and the students' vocabulary learning outcomes, H(2) = 7.775, p < 0.05. Pairwise comparisons were conducted to further examine the relationship between the identified groups and vocabulary learning outcomes. Pairwise comparisons showed significant differences in post-vocabulary test scores between students in Cluster 1 and Cluster 2 (p < 0.05). The students in Cluster 2 performed significantly better than students in Cluster 1. However, none of the other comparisons were significant after the Bonferroni correction (all p values >0.05).

Groups	Ν	Mean	SD	Mean Rank	X ²	р
Cluster 1	19	68.21	18.59	16.53	7.775	0.021*
Cluster 2	20	86.80	28.43	27.98		
Cluster 3	5	81.00	24.83	23.30		

Table 2. Results of Kruskal Wallis tests of post-vocabulary learning outcomes

Conclusions

The findings of this study were summarised into three aspects. First, this case study showed that primary students' SRVL behaviours leveraged by the mobile app with a self-regulation scheme could be clustered into three groups. Second, the similarities and differences of SRVL behavioural patterns using process-mining techniques in the three clusters were discussed. Compared to students in Cluster 1 and Cluster 3, students in Cluster 2 adopted more comprehensive SRVL learning trajectories using the mobile app with a self-regulation scheme. Third, a significant association between the identified three clusters and the students' vocabulary learning outcomes was observed, and students in Cluster 2 outperformed the other students.

This research was not without limitations. First, the study duration was short. Future studies should be conducted to understand students' SRVL behaviours over a longer period of time. Second, this study mainly relied on quantitative data. Future studies should include qualitative data (e.g., interviews) to explore the detailed characteristics of students' SRVL behaviours.

Acknowledgements

The study was funded by the General Research Fund (Ref. 18611019), Research Grants Council, University Grant Committee, Hong Kong.

References

- Dinh, D. T., Fujinami, T., & Huynh, V. N. (2019). Estimating the optimal number of clusters in categorical data clustering by silhouette coefficient. In J. Chen, T. Hashimoto, X. Tang & J. Wu (Eds.), *Knowledge and systems sciences* (pp. 1-17). Springer.
- Lehmann, T., Hähnlein, I., & Ifenthaler, D. (2014). Cognitive, metacognitive and motivational perspectives on preflection in self-regulated online learning. *Computers in Human Behavior*, *32*, 313-323.
- Li, S., Chen, G., Xing, W., Zheng, J., & Xie, C. (2020). Longitudinal clustering of students' self-regulated learning behaviors in engineering design. *Computers & Education*, 153. https://doi.org/10.1016/j.compedu.2020.103899

- Jang, E. E., Lajoie, S. P., Wagner, M., Xu, Z., Poitras, E., & Naismith, L. (2017). Person-oriented approaches to profiling learners in technology-rich learning environments for ecological learner modeling. *Journal of Educational Computing Research*, 55(4), 552-597. https://doi.org/10.1177/0735633116678995
- Matcha, W., Gašević, D., Uzir, N. A. A., Jovanović, J., & Pardo, A. (2019, March). Analytics of learning strategies: Associations with academic performance and feedback. In D. Azcona & R. Chung (Eds.), *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 461-470). Association for Computing Machinery.
- Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field: when measurement and intervention come hand in hand. Scandinavian Journal of Educational Research, 60(6), 723-735. https://doi.org/10.1080/00313831.2015.1066436
- Peeters, W., Saqr, M., & Viberg, O. (2020). Applying learning analytics to map students' self-regulated learning tactics in an academic writing course. In H. -J. So, M. M. Rodrigo, J. Mason, A. Mitrovic (Eds.), *Proceedings of the 28th International Conference on Computers in Education* (Vol. 1, pp. 245-254). Asia-Pacific Society for Computers in Education.
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451-502). Elsevier.
- Saint, J., Gašević, D., Matcha, W., Uzir, N. A. A., & Pardo, A. (2020, March). Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning. In A. Ahmad & I. Jivet (Eds.), *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 402-411). Association for Computing Machinery.
- Saks, K., & Leijen, Ä. (2019). The efficiency of prompts when supporting learner use of cognitive and metacognitive strategies. *Computer Assisted Language Learning*, 32(1-2), 1-16. https://doi.org/10.1080/09588221.2018.1459729
- Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word-class knowledge. *Language Testing*, 16(2), 189-216.
- Yang, K. -H., Chu, H. -C., & Chiang, L. -Y. (2018). Effects of a progressive prompting-based educational game on second graders' mathematics learning performance and behavioral patterns. *Journal of Educational Technology & Society*, 21(2), 322-334.
- Yang, Y., Wen, Y., & Song, Y. (in press). A systematic review of technology-enhanced self-regulated language learning. *Educational Technology & Society*, 26(1), 31-44.
- Zheng, L., Li, X., & Chen, F. (2018). Effects of a mobile self-regulated learning approach on students' learning achievements and self-regulated learning skills. *Innovations in Education and Teaching International, 55*(6), 616-624. https://doi.org/10.1080/14703297.2016.1259080
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *4*1(2), 64-70. https://doi.org/10.1207/s15430421tip4102_2

Yurika Ito

Waseda University, Tokyo

yurika.ito@akane.waseda.jp

A two-year investigation of a Facebook community for supporting language teachers using technology: Possibilities and challenges

Bio data



Yurika Ito is currently a PhD candidate and Research Associate at Waseda University, Tokyo, Japan. Her research interests include computer-assisted language learning (CALL), language teacher education, and the use of social media in language teaching and learning.

Abstract

The importance of educating language teachers to develop their competence in using technology for teaching purposes has been well established in the literature, but the reality is that many teachers around the world have been forced to educate themselves with little or no formal training. The Internet has opened up opportunities for individuals from all over the world to learn. With the use of Social Networking Sites (SNSs), language teachers are now able to connect with other like-minded teachers and support each other regardless of their geographic location and/or financial situation. Despite the increased number of online language teacher communities forming on these platforms, much remains unknown when it comes to how language teachers actually make use of such communities to learn about technology. Hence, the current mixed-methods longitudinal study investigated the role of online communities on SNSs for supporting language teachers who are using technology in their classes. The main data collection methods employed in the study were: (1) a content analysis of the posts shared in a Facebook community which consists of over 1000 online community members during the two-year observation period (i.e., October 2018 - September 2020), a questionnaire to understand the members' backgrounds and experiences in the community, and semi-structured interviews with the questionnaire respondents to obtain an in-depth understanding of their views about being a community member. The findings will be discussed in terms of the benefits and issues of using Facebook communities as a source of professional and emotional support.

Conference paper

Introduction

For language teachers to be able to teach using "smart" CALL technologies, they essentially need to become a "smart" CALL teacher. Smart CALL teachers are able to select appropriate CALL technologies and effectively integrate them into their classes to create a smart learning environment for their students. Integrating CALL technologies into language classrooms, however, is not an easy task. Merely providing language teachers with these technologies will not automatically result in successful integration (Hubbard & Levy, 2006). They need to be aware of the different technologies available,

learn how to use them, and learn how to implement them into their classrooms. What is more, with the constant advancements of technologies, they need to continuously enhance their skills and knowledge in technology throughout their entire teaching careers (Son, 2018). Although language teachers should be educated in how to use technology for language teaching purposes, in reality, not many of them are able to access formal CALL training programmes, often due to time and financial constraints and a lack of effective CALL training programmes (e.g., Kessler, 2007; Stockwell, 2009). Receiving little or no support from their colleagues, institutions, or government, they may have no choice but to learn on their own through informal means.

With the emergence of numerous online language teacher communities on SNSs over the past decade, SNS platforms seem to have the potential to serve as an alternative professional learning context for language teachers using technology. Previous studies (e.g., Carpenter & Krutka, 2015; Kelly & Antonio, 2016) have found that teachers can easily create new social relationships with other teachers who are teaching in similar situations on these platforms. They are able to join online teacher communities on SNSs and discuss matters concerning their views and queries about using technology for language teaching purposes. These communities could potentially help language teachers find new online resources, teaching ideas, and solutions to the problems they encounter when using technology. To date, studies on online teacher communities have predominantly focused on non-language teaching contexts (cf. Wesley, 2013), and there have been even fewer studies examining how online communities can specifically support language teachers using technology. Therefore, the current study aims to shed light on the realities of how language teachers are utilising these communities to learn about technology for teaching purposes. Although the current study was largely exploratory, it was guided by the following main research questions:

- 1) What is happening in a technology-focused language teacher community on Facebook?
- 2) What are the possibilities and challenges of being in such a community?

Methods

The current study is part of a larger research project, in which the role of online communities on SNSs for supporting language teachers who are learning about technology in language teaching and learning, is being investigated. Relevant studies in the past have often employed self-reporting data collection methods such as questionnaires and interviews with community members (e.g., Bissessar, 2014; Carpenter & Krutka, 2015), observations of posts shared in online communities on SNSs (e.g., Goodyear et al., 2019; Kelly & Antonio, 2016), and a combination of these methods (e.g., Curwood & Biddolph, 2018; Yildirim, 2018). Hence, for the current study, a combination of these methods were employed. Specifically, a technology-focused online teacher community on Facebook, which consisted of mostly language teachers in Japan, was observed for a period of two years from October 2018 to September 2020 to create a taxonomy of different types of posts shared in the community, and a questionnaire and semi-structured interviews with community members were administered to examine the community from a member's point of view.

Using content analysis as the main data analysis technique, the collected online posts and interview data were categorised and coded. The other observational data including the number of new members, number of posts, shares, likes, response rate for each post, and questionnaire responses were summarised using descriptive statistics.

Ethical considerations

Researching online communities on SNSs poses a number of ethical challenges. Considering the contentious ethical issues associated with investigating closed communities on Facebook (Kelly & Antonio, 2016), the current study focused on a public

language teacher community on Facebook which is open to the general public. Although ideally, it would have been preferred to seek informed consent from all community members, realistically speaking, seeking informed consent from all the members of a large online community with over 1000 members was not possible. Thus, since the community was a public group and the discussion topics were mostly related to technology and not particularly sensitive (McKee & Porter, 2009), informed consent was not taken. On the other hand, informed consent was received for the questionnaire and interviews.

Overview of the findings

Various types of online posts were identified in the observed technology-focused language teacher community on Facebook during the two-year observation period. The most frequent types of posts were about sharing upcoming events (e.g., conferences, webinars, social gatherings), courses, reading materials (e.g., books, research articles), and videos (e.g., Youtube videos), thereby indicating that the community is offering various learning opportunities to its members. Moreover, the online language teacher community was providing some of its members with professional and emotional support. Online community members were asking questions related to technology and teaching in the community. They were also using the online communities to connect to other teachers teaching in similar contexts and discussing teaching and non-teaching related matters. Although participating in such a community may bring various benefits, the challenges should not be overlooked. From the data sources, language teachers indicated that the online communities were not always a positive environment. Several participants reported that they had personally received rude and negative comments when participating in the discussions, and a few also reported that they had occasionally witnessed online arguments in the communities. Another concern identified from the findings relates to the quality of the posts. Several participants indicated that they were concerned about the credibility and reliability of posts and comments shared in the communities. Finally, another concern is the blurring of lines between work and private time. Although teachers can access these communities at any time of the day, having a lack of clear boundaries between work and private life may lead to negative consequences. Overall, the results of the analysis seem to indicate that technologyfocused language teacher communities on Facebook can be a double-edged sword for language teachers using technology.

Future directions

Based on the findings, it is apparent that further research is required to fully understand how online teacher communities on SNSs can support language teachers using technology. Since the current study only focused on language teachers who were already using online teacher communities on SNSs for professional learning purposes, the next step of the research project is to go beyond the original scope and include the views of other types of language teachers. It is worth expanding the sample to: (1) language teachers who members of online teacher communities on SNSs in the past but later quit, (2) those who are currently not using SNSs for professional purposes but are interested in using them for professional learning purposes, and (3) those who are resistant towards the use of SNSs for professional learning purposes.

Final remarks: Connecting the findings to the conference theme

According to Colpaert and Stockwell (2022), the "smart" in SMART CALL consists of three key elements, namely "personalization," "contextualization," and "socialization" (p. 4). They postulate that an effective learning environment possesses these three main elements. Based on this assumption, the findings from the current study seem to suggest that technology-focused language teacher communities on Facebook are providing language teachers with a "smart" learning environment: Language teachers who joined the observed online language teacher community on Facebook were exploring new ideas and obtaining resources and information about conferences, seminars and other relevant

events, which could potentially initiate and lead to the enhancement of their professional learning, especially with regards to learning about technology. The community seems to support personalised learning as they enable language teachers to find topics that they are personally interested in, which, in turn, could potentially help them in their own teaching context. Moreover, in the community, language teachers were socialising with one another by engaging in informal discussions with other community members and participating in the Friday night social gatherings, conferences, and webinars. They were also occasionally reaching ideas. It seems that these online language teacher communities on Facebook are being used as a platform for teachers to socialise with other teachers interested in learning about technology for teaching purposes. Although these online teacher communities are not without flaws, they appear to be offering various benefits to language teachers, particularly those who have limited professional learning resources and collegial support.

References

Bissessar, C. S. (2014). Facebook as an informal teacher professional development tool. <i>Australian Journal of Teacher Education</i> , 39(2), 121-135. http://doi.org/10.14221/aite.2014v39n2.9
Carpenter, J. P., & Krutka, D. G. (2015). Engagement through microblogging: Educator professional development via Twitter. <i>Professional Development in Education</i> , 41(4), 707-728. https://doi.org/10.1080/19415257.2014.939294
Colpaert, J., & Stockwell, G. (2022). Smart CALL: The concept. In J. Colpaert, & G. Stockwell (Eds.), <i>Smart CALL: Personalization, contextualization, & socialization</i> (pp. 1–6). London: Castledown Publishers. https://doi.org/10.29140/9781914291012-1
Curwood, J. S., & Biddolph, C. (2017). Understanding Twitter as a networked field site. In M. Knobel & C. Lankshear (Eds.). <i>Researching new literacies: Design, theory,</i> <i>and data in sociocultural investigation</i> (pp. 81-103). Peter Lang.
Goodyear, V. A., Parker, M., & Casey, A. (2019). Social media and teacher professional learning communities. <i>Physical Education and Sport Pedagogy</i> , 24(5), 421-433, https://doi.org/10.1080/17408989.2019.1617263
Hubbard, P., & Levy, M. (2006). The scope of CALL education. In P. Hubbard & M. Levy (Eds.). <i>Teacher education in CALL</i> (pp. 3-20). John Benjamin Publishing.
Kelly, N., & Antonio, A. (2016). Teacher peer support in social network sites. <i>Teaching</i> and Teacher Education, 56. 138-149. http://doi.org/10.1016/j.tate.2016.02.007
Kessler, G. (2007). Formal and informal CALL preparation and teacher attitude toward technology. <i>Computer Assisted Language Learning</i> , 20(2), 173-188. https://doi.org/10.1080/09588220701331394
McKee, H. A., & Porter, J. E. (2009). <i>The ethics of Internet research: A historical, case-based process</i> . Peter Lang.
Son, JB. (2018). Teacher development in technology-enhanced language teaching. Palgrave Macmillan.
Stockwell, G. (2009). Teacher education in CALL: Teaching teachers to educate themselves. <i>Innovation in Language Learning and Teaching</i> , 3(1), 99-112. https://doi.org/10.1080/17501220802655524
Yildirim, I. (2019). Using Facebook groups to support teachers' professional development. <i>Technology, Pedagogy and Education</i> , 28(5), 589-609. https://doi.org/10.1080/1475939X.2019.1686714
Wesley, P. M. (2013). Investigating the community of practice of world language

Wesley, P. M. (2013). Investigating the community of practice of world language educators. *Journal of Teacher Education*, 64(4), 305-318. http://doi.org/10.1177/0022487113489032

Proceedings of the XXIst International CALL Research Conference

215