ReCALL (2022), 34: 2, 127–129
doi:10.1017/S0958344022000040

CAMBRIDGE UNIVERSITY PRESS

EDITORIAL

Editorial

Shona Whyte

May is here and with it comes a new issue of *ReCALL* for 2022. This time, we have seven papers covering three broad topics: the first involves automatic L2 proficiency assessment, the second, mobile-assisted language learning (MALL), and the third, L2 vocabulary acquisition, particularly the role of audiovisual materials. Several papers draw on Paivio's (1971) dual processing theory, according to which "knowledge representation in verbal and visual modes may facilitate processing and therefore aid understanding and retention of knowledge more effectively than representations depending on a single mode" (Sato, Lai & Burden). There is also some overlap between the MALL and vocabulary studies, since the MALL meta-analysis by Burston and Giannakou reveals that by far the most common MALL learning objective is, in fact, lexical learning. Similarly, Lin's paper on a new web-based app focusing on formulaic expressions in YouTube videos sits at the intersection of MALL and learning of vocabulary and phraseology. As you will read, the studies include a range of methodologies and research designs, from meta-analysis (Burston & Giannakou; Yu & Trainin), to survey (Puebla, Fievet, Tsopanidi & Clahsen), experimental study (Dziemianko; Sato *et al.*) and computer modelling (Gaillat *et al.*), through to research and development (Lin).

Our first paper shows that collaborative research between universities in Paris and Galway is making headway in the complex area of automatic L2 proficiency assessment by developing AI systems to analyse learners' writing samples and assign them to appropriate CEFR proficiency levels. The research by **Thomas Gaillat, Andrew Simpkin, Nicolas Ballier, Bernardo Stearns, Annanda Sousa, Manon Bouyé** and **Manel Zarrouk** focuses on machine learning from a large corpus of Cambridge and Education First essays, using linguistic microsystems constructed around L2 functions such as modals of obligation, expressions of time, and proforms, for instance, in addition to more traditional measures of complexity involving lexis, syntax, semantics, and discourse features. After training on some 12,500 English L2 texts written by around 1,500 L1 French and Spanish examinees, which had been assigned to one of the six CEFR levels by human raters, the AI system reached 82% accuracy in identifying writers' proficiency levels. It also identified specific microsystems associated with learners at level A (nominals, modals of obligation, duration, quantification), level B (quantifiers and determiners), and level C (proforms and *should/will*). External validation for the model was less successful, however: only 51% of texts from the ASAG corpus (a different set of graded short answers) were correctly identified using logistic regression, rising to 59% with a more sophisticated elastic net method.

Moving on to the MALL papers, **Jack Burston** and **Konstantinos Giannakou** report on an extensive meta-analysis of a large number of studies published over the past quarter century in established CALL and education technology journals as well as in graduate theses. Their work focuses on research on learning outcomes and shows that around half the studies reporting learning effects were conducted at university level, most in interventions lasting 8–14 weeks, and most frequently in Asia, the Gulf States, and the US. Of the studies reviewed here, 95% had English as the target language, and, as noted, by far the most common learning objective

Cite this article: Whyte, S. (2022). Editorial. *ReCALL* 34(2): 127–129. https://doi.org/10.1017/S0958344022000040

© The Author(s), 2022. Published by Cambridge University Press on behalf of European Association for Computer Assisted Language Learning.

https://doi.org/10.1017/S0958344022000040 Published online by Cambridge University Press

was vocabulary. Statistical analyses using Hedges's *g* revealed substantial effect sizes (0.72 for studies with between-groups design, 1.16 for within-groups) but also high heterogeneity values (90% and 94% respectively). The authors interpret the findings to mean that MALL interventions are effective, but perhaps too different from one another to make it easy to aggregate results. They call for ongoing monitoring of MALL research, with perhaps a particular eye to work appearing in less traditional outlets.

Cecilia Puebla, Tiphaine Fievet, Marilena Tsopanidi and **Harald Clahsen** at the University of Potsdam also focus on MALL, but this time in older adults, focusing specifically on their uptake of language learning apps and factors affecting this engagement. This mixed-methods study employed an online questionnaire with some 200 older German participants (the majority aged mid-60s to mid-70s) and interviewed a separate cohort of around 20 with similar profiles. The group was generally well educated (more than half had attended university) and around 80% were retired. Regarding technology use, almost three quarters owned a smartphone, two thirds a laptop, and around half the participants owned a desktop computer; most reported high digital literacy and few problems with technology. However, of the sizeable proportion of respondents who reported using language learning apps (around one third of the group), less than half were satisfied with their experience. The study found links between digital fluency and appreciation of MALL apps, and negative correlations with age and desktop computer use. In a detailed discussion section, the researchers note that these older learners often prefer computers for cognitively demanding activities, and that requirements for personalisation, connection, and authenticity associated with both mobile learning and geragogy (learning theory for older groups) did not seem to be met for older learners in the currently available MALL offerings, thus disincentivising this population from taking up language learning opportunities in this form.

The next paper reports on the author's initiative to design an appropriate language learning experience that capitalises on learners' existing informal online language practices. **Phoebe Lin**'s paper describes the functions and development of IdiomsTube, a web-based app for L2 learning of formulaic expressions in English from learner-selected, English-captioned YouTube videos. Launched in 2018, the app assesses the level of a given video, generates learning tasks, and suggests other relevant videos. After reviewing the literature on formulaic expressions and the difficulties they present L2 learners, the paper explains how videos are classified using BNC/COCA word frequency bands, then presents the three types of learning tasks. Pre-learning activities involve a glossary with dictionary lookup using reliable online sources; consolidation activities focus on meaning (gapfill), form (hangman), and pronunciation (self-recording) exercises; revision tasks include flashcards for systematic review. There is also a teacher interface to allow integration with other classroom activities.

The following two papers involve classroom computer-based learning of vocabulary. **Takeshi Sato, Yuda Lai** and **Tyler Burden**'s study draws on a number of different theoretical foundations to compare the effectiveness of static versus dynamic visual aids: Paivio's dual coding theory, image schemata as conceived in cognitive linguistics by Langacker, for example, and style of processing, whereby learners are classified as "imagers" or "verbalisers." The researchers also include L1 transfer as an area of interest in their experimental study of L2 English learning of the three spatial prepositions *on*, *over* and *above* by speakers of L1 Taiwanese, which makes a two-way distinction in this semantic area, and L1 Japanese, which makes no distinction at all. The experiment involved teaching physical and metaphorical senses of the three prepositions to Japanese and Taiwanese learners with the support of either static images or animations in a pre-, post- and retention-test design, using a 40-item cloze passage where learners were asked to supply the appropriate target preposition. Results showed that imagers generally learned better, but there was no difference between static and dynamic images. Regarding L1 effects, however, a slight advantage in immediate uptake was shown for the Taiwanese participants who were shown dynamic images, and a similar advantage for Japanese speakers appeared on the retention test two weeks later.

https://doi.org/10.1017/S0958344022000040 Published online by Cambridge University Press

Anna Dziemianko's study also investigated the effect of graphic illustrations on vocabulary learning, this time in online dictionary entries. Some 200 Polish students with B2 English learned 15 items in one of four conditions: with colour photos, greyscale photos, line drawings or no illustration. The researcher measured accuracy and time required for comprehension, then immediate and delayed retention. She found that although all forms of illustration led to better learning than no illustrations, the group that was shown line drawings learned more quickly and retained more items in both the test following initial exposure and two weeks later. Items with colour images were also quickly picked up but less well retained, while greyscale images resulted in longer processing times and offered no better delayed retention than dictionary entries without illustrations.

The last paper in this issue is **Aiqing Yu and Guy Trainin**'s meta-analysis of technology-assisted L2 vocabulary learning, which examines a 12-year window (2006–2017), looking at overall learning outcomes and also the specific role played by type of instruction, assessment, L1–L2 distance, learner age, and technology type. The researchers analysed 34 studies with 49 effect sizes and found a moderate overall effect size (Cohen's $d = 0.64$). Looking in more detail, they discovered that university learners outperformed school pupils, and closer L1–L2 pairings (e.g. two Indo-European languages versus one Indo-European and one non-Indo-European) resulted in better vocabulary learning. With respect to technology, interventions using mobile devices produced greater learning effects than computer-based studies, and regarding instructional objectives, interventions aimed at incidental vocabulary learning were more effective than those focusing on deliberate learning. Finally, although the assessment of receptive learning (e.g. recognition of a word in a multiple-choice question) showed a larger effect size than productive learning (e.g. sentence completion), this difference was not statistically significant.

So those are the studies you can read in the May issue. As you might imagine, both of the meta-analyses that appear here, like many in this genre, call for greater methodological rigour and more transparent reporting in primary studies. We certainly hope that the other experimental studies in this issue may inspire further replication and indeed meta-analysis – *ReCALL* authors and editors are certainly doing their best.

https://doi.org/10.1017/S0958344022000040 Published online by Cambridge University Press