Bilingualism: Language and Cognition

cambridge.org/bil

Digital Language Learning (DLL): Insights from Behavior, Cognition, and the Brain

Ping Li[1] and Yu-Ju Lan[2]

[1]The Hong Kong Polytechnic University, Hong Kong and [2]National Taiwan Normal University, Taipei

Keynote Article

Cite this article: Li P, Lan Y-J (2022). Digital Language Learning (DLL): Insights from Behavior, Cognition, and the Brain. *Bilingualism: Language and Cognition* **25**, 361–378. https://doi.org/10.1017/S1366728921000353

Received: 14 June 2021
Revised: 24 June 2021
Accepted: 24 June 2021
First published online: 13 August 2021

Keywords:
digital language learning; second language acquisition; bilingual representation; social learning; neurocognition of language

Address for Correspondence:
Ping Li, Department of Chinese and Bilingual Studies, Faculty of Humanities, The Hong Kong Polytechnic University, Hong Kong SAR, China.
Email: pi2li@polyu.edu.hk

Abstract

How can we leverage digital technologies to enhance language learning and bilingual representation? In this digital era, our theories and practices for the learning and teaching of second languages (L2) have lagged behind the pace of scientific advances and technological innovations. Here we outline the approach of digital language learning (DLL) for L2 acquisition and representation, and provide a theoretical synthesis and analytical framework regarding DLL's current and future promises. Theoretically, DLL provides a forum for understanding differences between child language and adult L2 learning, and the effects of learning context and learner characteristics. Practically, findings from learner behaviors, cognitive and affective processing, and brain correlates can inform DLL-based language pedagogies. Because of its highly interdisciplinary nature, DLL can serve as an approach to integrate cognitive, social, affective, and neural dimensions of L2 learning with new and emerging technologies including VR, AI, and big data analytics.

1. Introduction

Our society today is faced with significant challenges, one of which is the lack of effective communication through multiple languages. The challenge is further exacerbated by the outbreak of the Covid-19 pandemic that requires people to practice 'social distancing' and avoid 'social interaction.' Social distancing is fundamentally against human nature, and the prolonged practice has created not only economic hardships and cognitive disturbances, but also difficulties in language learning for both children and adults.[1] Thanks in large part to the pervasive use of digital technologies, we have dealt with some of the difficulties under the pandemic, from video conferencing to online teaching to virtual gathering. In the last decade, digital technologies have also developed alongside advances in artificial intelligence (AI) and big data analytics. These developments have changed human behavior in all aspects of our lives including how we learn a new language. Digital language learning (DLL) has emerged against this backdrop both as an educational practice and as a field of scientific study.

DLL can be used broadly to refer to digital technology-based or technology-enhanced language learning platforms or tools, or the practices of learning using such platforms or tools. In this paper, we use DLL in this broad sense to reflect the new developments through technology-driven methodologies, with second language (L2) learning as our focus. Although a DLL approach covers similar techniques encompassed by computer-assisted language learning (CALL), DLL focuses on more recent tools and platforms enabled by the latest developments in digital technologies such as mobile computing, virtual reality (from desktop 3D to augmented/mixed reality), and digital games, attempting to explore the potential of technologies for cultivating self-directed, exploratory, and autonomous learning.

Riding on the tide of rapidly developing digital technologies, L2 learners and teachers have delved into DLL and its applications. Indeed, DLL-based L2 platforms and tools have emerged so quickly in the past decade that we can no longer count them by our fingers. At the same time, however, it is unclear whether some of the commercial products (e.g., *Babbel, Duolingo, Rosetta Stone*) are always validated scientifically or empirically (see van Deusen-Scholl, 2015 for discussion). It is also unclear how we might go about assessing these tools against each other and against their bold commercial claims about their effectiveness, when no randomized control studies could be performed (which is a problem with some reports out there, e.g., Vesselinov & Grego, 2012, 2016). Further, significant gaps exist between the DLL tools that the tech companies develop and the needs that learners and instructors have. It is clear, though, that the industry does not always have in mind learner-specific characteristics or the assessment of learning success, as will be discussed in this article. Technology developers are mostly interested in making their products available (and gaining profits), whereas

[1]A recent post illustrates this point clearly: https://npjscilearncommunity.nature.com/posts/social-interaction-is-good-for-learning-a-new-language-social-isolation-is-not

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

CAMBRIDGE UNIVERSITY PRESS

educators are interested in using the technologies to enhance learning outcomes; unfortunately, these two do not always match. Such gaps are further complicated by the fact that even educators/instructors may not necessarily know what environmental and learner characteristics are relevant and critical without in-depth research efforts. Thus, insights from scientific studies of behavior, cognition, and the brain would be crucial.

The societal challenges, the advances in digital technologies, the gaps between DLL tools and their fit to learner characteristics, and the impacts of DLL on brain and behavior, form the bases of our discussion in this article. The purpose of this article is not to provide a comprehensive review of the literature; many such reviews already exist as discussed below, including special volumes (e.g., Chapelle & Sauro, 2017; Levy & Stockwell, 2006). Our goal in this article is to provide a theoretical synthesis and analytical framework with respect to DLL's current promises, theoretical and pedagogical implications, and future directions.

2. CALL in the past and DLL in the new era

To learn a new language in addition to one's first language (L1) is always challenging. It takes time, effort, attention, motivation, and sustained involvement. The ability to use a language for communication and social interaction is a critical competence needed by everyone in the 21st century. Technology has played a significant role in helping today's learners to acquire a language. In this article, we intend to examine DLL in the views of a wide range of methods and platforms enabled by new digital technologies such as mobile computing, VR, and digital games. CALL has dominated the field for over 30 years since computers became popular (see Otto, 2017 for a general overview of the history of technology and L2 learning). Many of the methods used by earlier CALL are still widely adopted as the standard methods today (e.g., gap-filling/cloze tests, multiple choices, flashcards, and sentence reordering, both in L2 classrooms and on the web), but fundamental differences exist between the earlier CALL efforts and today's massively interactive, web-based, app-based, and mobile-enabled DLL methods (see Presson, Davy, & MacWhinney, 2013 for an earlier argument in this regard).

Shifts in the use of technology for language learning and teaching, as with the general trends in education, can be observed in terms of different emphases and focuses of the time based on different theoretical foundations, technological development, and educational paradigms. As described by Warschauer (2004), between the 1970s and the 1980s, the behaviorist paradigm had dominated language learning and computer-assisted teaching – that is, the entire CALL field; during this period, the computer-learner were treated in a stimulus-response relationship due to behaviorism, and drill-and-practice remained the main method. The cognitive approach rejected behaviorism for language learning in the 1980s and the 1990s, although the actual paradigm shift from behaviorism to cognitivism occurred two decades earlier (see Gardner, 1984). During this period communicative exercises were emphasized, and fluency, rather than language analyses and grammar, was the major focus of language teaching. CALL software and language games also began to flourish during this period. Next, in the 2000s, the authentic context of learning and social interaction was highlighted (see Otto, 2017) and social-cognitive dimensions of learning shed light on language education and research. These developments also grew alongside the increasing popularity of social media and multimedia technologies (e.g., videos that can incorporate text, graphics, audio, and animations; Mayer, 2005).

Based on Warschauer's (2004) perspective, Chun (2019) expanded the framework by adding to the focus of DLL in the 2010s seamless digital technologies, technologies that have extended language learning spaces and blurred the boundaries of formal and informal learning. Learning is no longer isolated from the environment; instead, it is embedded in the context in which authentic learning takes place. This development goes hand-in-hand with today's focus on e-learning, blended learning, and multimedia learning, aided significantly by ubiquitous computing, mobile apps, and wearable devices. Such technological advances have greatly promoted multimedia and multimodal learning in all subject areas, and in the last year due to the pandemic, the pace of development has been further accelerated.

With these paradigm shifts for language learning in the past decades, we predict that DLL in the third decade of the 21st century will further focus on new approaches. In particular, big data and AI are impacting every aspect of our lives and our society, from the environment (energy, climate, ecosystem, space) to human behavior (aging, health, education). AI technologies, such as machine learning, automatic speech recognition, and natural language processing (NLP), no doubt also have profound implications for education (Luan, Geczy, Lai, Gobert, Yang, Ogata, Baltes, Guerra, Li & Tsai, 2020). Language learning is no exception in this regard (see details in Section 5). We have seen an unprecedented increase in the integration of AI and language applications: for example, mobile apps with image recognition and NLP turn the real world into a language learning setting; automatic evaluation systems analyze the errors in L2 learners' writings (Al-Ahdal, 2020) and provide instant feedback on correct grammar and hints on best writing (e.g., the popular *Grammarly* software; see some earliest efforts discussed in Grosjean, 2019, Chapter 7); the combination of VR and intelligent agents creates immersive and authentic contexts allowing language learners to have social interaction in real-life like situations (e.g., Nicolaidou, Pissas & Boglou, 2021); and virtual agents through interactive dialogues can enhance learners' language performance (e.g., Graesser, Chipman, Haynes & Olney, 2005; Junaidi, Hamuddin, Julita, Rahman & Derin, 2020; Tai & Chen, 2020); these are only a few of the many examples in recent years.

To truly take advantage of the AI technologies, we must also make use of the big data readily available during language learning, along with the relevant data analytic tools. For example, in a smart learning environment, the entire learning process can be logged on a key-stroke or step-wise level, and the learner data can be automatically analyzed and visualized. Based on such analytic results, a personalized learning plan can be recommended and the learning materials that fit individual learning profiles can be appropriately provided (see 3.1; Kokoç, Akçapınar & Hasnine, 2021; Yang, Chen & Ogata, 2021). Better still, such personalized feedback can be provided in real time, providing instant information to allow learners to adjust their pace as they learn, to see their up-to-the-point achievements, weaknesses, and learning behavior patterns. For the learner, learning opportunities are available anywhere and anytime (Pikhart, 2020); for the educator and researcher, making use of the data generated in such environments would guide the design and implementation of precise and personalized education (Godwin-Jones, 2017; Lan, 2016; Yang, 2021).

While all these developments are exciting, we must caution that learning successes through DLL are not automatically guaranteed. As pointed out by Godwin-Jones (2019), the ability to

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

conduct self-regulated, self-directed, and self-reflective learning is essential to learners' language acquisition. Furthermore, learning outcomes obtained in a DLL environment cannot be precisely evaluated simply by traditional achievement tests. Multifaceted evidence should be leveraged to correctly evaluate the effects of DLL, which may include data-driven analyses of learning patterns and behaviors. Based on this consideration, it is also important to understand characteristics of learning (e.g., what make up a better or worse DLL environment) and individual differences of the learner (e.g., cognitive ability, language aptitude, and learning strategies). Finally, we have shown that learning a new language through innovative technologies brings about positive changes in the learner's brain structure and function (e.g., Legault, Fang, Lan & Li, 2019b) but so far we have only limited knowledge in this regard. To study how L1 vs. L2 neural representations emerge as a function of DLL will surely be a new exciting direction (see 4.4).

3. New developments in DLL: MALL, VR, and game-based language learning

In this section, we highlight recent developments in digital technologies and their applications in DLL – in particular, mobile-assisted language learning (MALL), virtual reality (VR), and digital game-based language learning (GBLL). Shadiev and Yang (2020) listed 19 different technologies that have been used for language learning and teaching, many of which are based on the latest digital technologies. There is a tendency to name all technology-enabled language learning as CALL, but we think this does not do justice to a field that is so rapidly developing and that rides on the successes of new emerging technologies. In our view, 'computer-assisted' methods are being replaced, both IN PRACTICE and IN THEORY by emerging technologies and fields of studies (e.g., multimedia learning, blended learning, situated/embodied learning, and social learning). Furthermore, a new industry has joined hands with educational technology in designing popular digital tools and platforms for language learning. Several highly commercialized products attract millions of users to learn new languages (e.g., *Babbel*, *Duolingo*, *Rosetta Stone*), although their scope, languages covered, and functionality vary widely. As discussed earlier, technological innovations can drive pedagogical paradigm shifts, and in the case of language education, shifts are occurring from the classroom-based, instruction-oriented, and teacher-centered approaches to student-centered teaching and learning, as in other areas of education. Covid-19 has brought a 'new normal', and digital technologies play an even more critical role now than ever before. DLL rides on this tide to move from the 'computer-assisted' ideas and methods to the massively socially connected, web-based, and app-based MALL, VR, GBLL tools and platforms, enabling contextualized and embodied language learning to occur in the real or simulated real world.

3.1. Mobile-Assisted Language Learning (MALL)

The popularity of MALL has increased dramatically as mobile devices such as smartphones and tablets become indispensable in our daily lives. Ubiquitous as they are, mobile technologies now provide convenient platforms to support L2 learning anytime and anywhere, overcoming the limitations imposed by physical classrooms. Importantly, MALL allows the learner to acquire a new language through real-life exploration, effectively turning the real world into a learning context. Situated learning (Anderson, Reder & Simon, 1996; Dede, 2009; Dede, Jacobson & Richards, 2017) is a concept referring to learning taking place in real-world or real-world like (simulated) situations, which can be implemented on mobile devices or through immersive technologies (see 3.2). Such situated MALL platforms can connect L2 learning with real-life events and contexts (Lin & Lin, 2019; Lin, Lin, Liu, Kou, Kulikova & Lin, 2020; Shadiev, Hwang & Huang, 2017), and many MALL applications also involve game playing which promotes language learning through self-exploratory and knowledge construction processes. An additional new direction has been to integrate MALL with data-driven and AI-inspired methodologies, such as automatic speech recognition, NLP, and image recognition, resulting in many new web-based tools or apps (e.g., Chen, Yang & Lai, 2020; Shadiev, Zhang, Wu & Huang, 2020).

Drawing from Kearney, Schuck, Burden and Aubusson's (2012) framework for mobile learning, Lai and Zheng (2018) identified three key features that make MALL significant for L2 learning: personalization, authenticity, and connectivity. By surveying many college students with follow-up interviews, the authors found that the students used MALL mostly for their personal learning purposes, and less for authentic language learning or social connection. In a more recent review, Tu, Zou and Zhang (2020) expanded on these features to include portability, real-time interaction, and situated learning, but also reviewed the negative aspects of MALL such as limited screen space and users' short attention span for learning. Some commercial products such as *Google Translate* provide situated learning through instant phone camera translations while *Instagram* and *WhatsApp* enable social networking groups to learn L2 and chat with native speakers on the phone. Tu et al. also articulated an evaluation framework for MALL apps designed for vocabulary learning in terms of factors such as content quality, multimodal presentation, engagement, and usability. While most MALL applications are designed for young adults, Puebla, Fievet, Tsopanidi and Clahsen (2021) conducted a web-based survey with over 200 participants and further in-depth interviews to see whether older adults are open to using MALL for learning L2. The authors found that older adults, unlike younger generations, are more resistant to adopting MALL applications, mainly because they dislike personal interactions that are not face-to-face. MALL and DLL in general have so far been focused on young adults or college students, and their application and use for older adults thus require further examination (see also Wang & Christiansen, 2019 who tested a population with a mean age of 51, which may be too young to count as 'older adults').

It has become popular for MALL applications to use QR codes attached to real objects to enable mobile phones to display L2 sounds and labels (e.g., Chinese characters). Liu, Chen and Hwang (2018) developed such a context-aware system for improving L2 English learners' listening comprehension. It allowed learners to scan QR codes attached to exercise machines in a fitness center to learn exercise-related vocabulary collaboratively. Other than vocabulary, higher-level language skills, such as conversational interactions and writing, can also benefit from mobile technology-based language tasks (e.g., Gharehblagh & Nasri, 2020; Lan & Lin, 2016). Previous work has found that students using MALL outperform their peers without MALL support; for example, in oral communication (Lan & Lin, 2016) and in English writing (Gharehblagh & Nasri, 2020). Even at the non-linguistic level, Lee, Lo and Chin (2021) showed that mobile technologies support the integration of multimodal information and

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

social interaction, which can trigger intercultural learning and increase multicultural awareness. Lomicka and Ducate (2021) also encouraged students to work with peers collaboratively, and through posts at Padlet, a social networking app, the students could share ideas and knowledge about culture and cultural experiences. Given the tight-knit relationships among sociocultural adaptation, intercultural learning, and L2 proficiency (Ward & Kennedy, 1996), MALL can play a significant role in enhancing both communicative competence and intercultural interaction.

A novel use of the MALL technology is to combine it with adaptive learning algorithms to enable the design of learning material to better fit student profiles (see Section 5 for further discussion). Sandberg, Maris and Hoogendoorn's (2014) adaptive model is an example of this: they weighted the 120 learning target words by different linguistic characteristics and derived a level of initial difficulty for each word, adjusting the level as student learning progressed. This way the MALL platform could create a dynamic student model that considers the learner's developing level of knowledge. Similarly, Stockwell (2007) described an intelligent vocabulary MALL system in which learners' access and performance information was tracked, and new exercises were automatically generated to fit the level of individual learners. Pandarova, Schmidt, Hartig, Boubekki, Jones and Brefeld (2019) further extended this approach to grammar learning, although not on MALL platforms. Theoretically, these approaches are also consistent with the 'input hypothesis' of second language theory (Krashen, 1988), according to which the target input for learning should be one level higher beyond the learner's current level of knowledge.

The evidence on MALL's overall effectiveness for L2 learning, as compared with other methods, remains mixed (Chen, Tseng & Hsiao, 2018; Loewen, Crowther, Isbell, Kim, Maloney, Miller & Rawal, 2019). For example, learning based on mobile apps, compared with CALL or in-person teaching, produced similar results for high school students (e.g., Peterson, 2010). It may be that high schoolers are a group of users with extremely high frequency of use of mobile phones for social networking, and this has negatively impacted their ability to make use of MALL effectively for content-based language learning. Recent behavioral and brain imaging data suggest that young people's excessive use of electronic devices including mobile phones may have adverse effects on scientific knowledge integration (Hsu, Clariana, Schloss & Li, 2019) and on Chinese literacy development (Tan & Xu, 2020). Another significant limitation of mobile devices on learning is their small screens, which may increase learners' cognitive load, especially when the processing of rich and multi-page information is necessary. In addition, older adults may find the screen's small size a particular weakness of MALL when they need face-to-face interaction (see Puebla et al., 2021). Finally, except a few recent studies most MALL applications remain limited to basic skills such as vocabulary learning (Lai & Zheng, 2018; Lin & Lin, 2019). Given this limitation, some authors (e.g., Hannibal Jensen, 2019; Presson et al., 2013; Sykes, 2017) called for the use of extended mobile technologies to include videos, social media, and Google maps to enhance not just vocabulary learning but also other communicative skills.

3.2. Virtual Reality (VR)

VR has emerged as an important technology for education in the last two decades because of its significant potential and impact on student learning in many educational contexts (see Li, Legault, Klippel & Zhao, 2020; Liu, Dede, Huang & Richards, 2017). The role of VR in student learning has received much attention, but its application in experimental studies of L2 learning has been more recent (see Legault, Zhao, Chi, Chen, Klippel & Li, 2019a; Li et al., 2020). The term VR can be used to cover a wide range of virtual environments and tools including: dynamic 3D displays projected on computer monitors (desktop or tablet virtual environments; VE); on large screens/walls in amphitheaters, rooms, or specialized cubicles outfitted for 3D images (e.g., CAVE systems); on head-mounted displays (HMD); through devices that show digital image enhancements ('augmented reality' or AR); and through a blend of virtual and real-world objects projected onto HMDs ('mixed reality' or MR). This broad range of VE, VR, AR, and MR vary in immersion (e.g., 360-degree views vs. limited wide-angle views), interactivity (extent of action and movement), social presence (whether there is feeling of being there), and ultimately realism (how realistically VR simulates the real world).

Broadly speaking, VR can be categorized into two types (Robertson, Card & Mackinlay, 1993): immersive VR (iVR) and non-immersive VR. Both types of VR aim at creating authentic (i.e., real-world like) environments to enable learning through active and self-exploratory discovery in the virtual environments (Dede, 2009). Among the many innovative applications of VR, social interaction through simulated immersion seems to be the most important for L2 learning (see 4.2-4.4). As argued by many theories of L2 acquisition, meaningful social interaction is one of the most significant processes that lead to the success of L2 acquisition (e.g., Ellis, 2019; Lantolf, 2006; Mackey, Abbuhl & Gass, 2012). Another significant advantage of VR, to educators and researchers alike, is its flexibility in designing learning contexts that can vary systematically in environmental characteristics (Casasanto & Jasmin, 2018). A real-world situation contains too many variables or noises that may confound a study, but VR enables modification and manipulation of virtual environments with rigorous control. In other words, VR provides both 'high ecological validity' and 'high experimental control', thereby lending researchers an excellent tool to study naturalistic events in the lab (Peeters, 2019).

Language learning in VR is contextualized and interaction-oriented. Like MALL, VR fulfills three essential components of successful L2 learning – that is, authentic contexts, learners' active involvement, and meaningful social interaction (see Lan, 2014; Legault et al., 2019a; Sadler, 2017 for reviews). Sadler (2017) provided a brief history of L2 applications of virtual worlds including platforms such as *Second Life*. Lan (2020a) suggested that current L2 applications of VR learning can be classified into five categories based on different pedagogical purposes: (1) expanding L2 learners' visual experience, (2) learning by operating or manipulating virtual objects, (3) learning by creation, (4) creating a joyful learning process, and (5) building a social network. First, the L2 learner can have enhanced visual experiences, particularly in immersive VR contexts. Such experiences may not only match with our visual experiences in the physical world, but also expand our experiences to transcend boundaries in time and space, such as attending a 17th-century drama play in Shakespeare's time, observing creatures under the sea, and walking in outer space, experiences not possible in the real world (Dede et al., 2017; Mohsen, 2016). For L2 learning, the student can easily be 'transported to' or immersed in regions where the target language is used, along with the relevant cultural artifacts and environmental characteristics. Second, they can manipulate or operate on the virtual objects as in the real environment, sometimes even with enhanced capabilities. For example, in Lan, Fang, Legault and

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

Li (2015) and Legault et al. (2019a), L2 learners can move spoons, cups, teapots, and other kitchenware in the VR Kitchen, experiencing the tactile and motoric aspects of the objects when learning the L2 words/labels. The learners could also walk along a path to see the animals in a VR Zoo. This type of tactile, sensory, motoric learning allows learners to contextualize the acquired labels – that is, to represent them in an embodied manner, closely matching with what the child does during L1 learning (see 4.1). Third, in addition to exploring the virtual worlds, sharing one's 3D creation before and during learning is also an innovative VR application (e.g., Yeh & Lan, 2018). Learning by creation strengthens learners' ownership and consequently promotes their learning autonomy (Lan, Hsiao, Fang & Chen, 2018). Fourth, getting immersed in VR worlds is a joyful experience for many users, allowing for learners' exploration of an unknown environment. In this regard, many studies have indicated that VR motivates students' positive attitudes towards learning (see Lan, 2015, 2020b for reviews). Fifth, VR enhances interpersonal interaction through multi-user platforms such as *Second Life*, allowing L2 learners to create a social community and interact with each other from around the globe. Such a social community also enables L2 learners to perform 'role playing' during language learning.

Verbal and non-verbal skills, from vocabulary to listening and from spoken conversation to interpersonal communication, can all be enhanced by VR, given VR's specific features of immersion, interactivity, and enabling of imagination and innovation (Lan, 2020a; Li et al., 2020). Further, Chen (2016) showed that virtual environments could enhance students' engagement and promote collaboration in communication. Even in studies that used basic desktop 3D virtual environments, researchers have found VR to help enhance learning outcomes. For example, Lan et al. (2015) constructed *Second Life* environments to train American students to learn Mandarin Chinese vocabulary. The authors showed that learning in *Second Life* needed only about half the number of exposures to attain the same level of accuracy as learning via computer-based picture-word paired associations; in addition, students showed faster acceleration of learning in the second phase of training. Such differences between VR learning and non-VR learning were further observed in immersive VR environments in Legault et al. (2019a).

As VR becomes more accessible and portable, more computational resources and tools are also available (e.g., Turbosquid 3D models and Unity development tools), which enables educators to develop real-life like environments more easily (e.g., garden, kitchen, library, MTR station, school, shopping mall, street, supermarket, and zoo). However, there remain a number of limitations of current VR-based applications for L2 learning: (a) sample sizes are small in most studies, limiting the generalizability of findings; (b) descriptive results, rather than statistically tested findings, are usually reported (see Wang, Lan, Tseng, Lin & Kao, 2020 for a discussion); (c) popular VR applications (and DLL tools in general) such as *House of Language VR* (Oculus Gear) remain limited in their scope of coverage and number of languages; (d) most of the popular VR headsets (e.g., *HTC Vive*) remain bulky, and may be unsuitable for younger users. These limitations, we believe, can be overcome in future large-scale studies with future technological developments that make VR more portable and easier to use.

3.3. Game-Based Language Learning (GBLL)

Young people are game lovers, especially the Millennials and the Generation Z who are the 'digital natives' growing up with smartphones, tablets, and online games. In the past decades, a significant amount of research interest has been directed to games for education (Mayer, 2016). Against this context, GBLL has become particularly popular in recent years. Although many of the CALL, MALL, and VR platforms discussed above are also game-based, researchers have treated GBLL as a separate methodological approach probably because games have had a longer tradition and wider usage than digital learning.[2] The idea here is that like other 'serious games', GBLL games are not just for fun or entertainment, but are explicitly structured with educational purposes and goals (e.g., learning L2 vocabulary). So far, most GBLL research has focused on learning English as an L2 (over 90% of the studies) and has used video gaming or immersive gaming platforms for single users and role-playing games for multi-users (for reviews see Hung, Yang, Hwang, Chu & Wang, 2018 and Reinhardt, 2017).

It is not yet clear how much gaming experience (e.g., frequency/amount of time, and proficiency in playing computer games) can affect the success of L2 learning. Hung et al. (2018) reviewed several studies that indicate a potential relationship between experience in digital games and the learner's L2 proficiency, particularly for male students (e.g., Smith, Li, Drobisz, Park, Kim & Smith, 2013; Sundqvist & Sylvén, 2012). However, the evidence so far is mixed regarding GBLL's effectiveness as compared with traditional methods of language learning (e.g., deHaan, Michael Reed & Kuwada, 2010; Sundqvist & Wikström, 2015). For example, Rachels and Rockinson-Szapkiw (2018) found that Spanish L2 learning using *Duolingo* and traditional teacher-student instruction did not make a difference; similarly, Loewen et al. (2019) found that students learning Turkish as L2 with *Duolingo* had shown limited gains, calling into question the overstated claims on *Duolingo's* efficacy (Vesselinov & Grego, 2012). Some meta-analyses (e.g., Cerezo, Baralt, Suh & Leow, 2014; Grgurović, Chapelle & Shelley, 2013) also indicated mixed results, with some showing an overall advantage of GBLL, while others showing similar performances with both GBLL and non-game based learning. Further, there may be individual differences, as Hung, Young and Lin (2015) showed that, for high-achieving students, gaming vs. non-gaming conditions did not make a difference, whereas for low-achieving students GBLL was more effective (see also Legault et al., 2019a for a similar pattern in VR vs. non-VR learning). Yu (2018) found that, for male more than female students, GBLL led to better English L2 learning than traditional approaches. The good news is that GBLL generally produces positive learning outcomes (e.g., Foomani & Hedayati, 2016; Sato, Murase & Burden, 2015; Shi, Luo & He, 2017), although this positive learning effect might be more evident for vocabulary than for other aspects (grammar, pronunciation, pragmatics; Hung et al., 2018; Tsai & Tsai, 2018; Zou, Huang, & Xie, 2019).

Acquah and Katz (2020) suggested six important features that make GBLL particularly appealing for language learning: ease of use, challenging, reward-and-feedback, control/autonomy, goal-directedness, and interactivity. Previous work has indicated that

[2]There are various terms used in the literature for language or non-language games, including gamification, serious games, digital learning games, action video games, multiplayer online role-playing games, and so on. For consistency, we use the term 'game-based language learning', or GBLL for short. As the majority of the work in this domain focuses on digital rather than non-digital games, we also do not use the longer acronyms of DGBLL (digital game-based language learning). See Hung et al. (2018; Figure 1) for an illustration of GBL, DGBL, and DGBLL.

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

games may activate the user's intrinsic motivation and provide learners with a sense of autonomy or control (e.g., Peterson, 2010). Like MALL and VR, GBLL engages attention, activates prior knowledge, and is often situated in real-life contexts. Acquah and Katz further pointed out that not all six features equally influence language learning; for example, challenging games can increase motivation, but not necessarily improve learning outcomes. Another feature not discussed by the authors is the adaptivity of games (as in MALL, see 3.1), and the extant evidence points to positive effects of adaptive educational games on learning achievement and engagement in general; see Liu, Moon, Kim, and Dai (2020) for a recent review.

Gaming itself is a social process that involves multiple users/parties. While many GBLL platforms have been developed for L2 learners to play on a 'one-on-one' basis, multiplayer environments – specifically, the 'massively multiplayer online role-playing games' (MMORPGs; Peterson, 2010) – have become important for language learning. Unlike single-user games, MMORPGs operate on connected networks, in real-time, and engage many people simultaneously in the same gaming environment or learning process (e.g., the most popular gaming platform *World of Warcraft*). According to Wimmer (2008), we should identify the important elements for 'dynamic interaction' in MMORPGs – including, at least, the learners, the environments, the objects in the environments, and the results of interactions among these elements. Peterson (2016) further extended these to include other features of MMORPGs: large number of users, use of personal avatars, real-time interaction, immersion in virtual worlds, game-embedded quests, and extensive user-created contents, which all may be highly relevant to language learning. From a cognitive perspective, unlike other GBLL tools, MMORPG games are particularly facilitative to L2 production, because the learner needs to develop a communicative ability by holding dialogues with other players in the language of the game (e.g., Reinders & Wattana, 2014; Suh, Kim & Kim, 2010). From a sociocultural perspective, MMORPGs provide learning environments that are conducive to socialization through language use, and help to develop a positive learner attitude (Peterson, 2016).

4. How does DLL matter? Insights from multiple dimensions

The new developments in DLL as discussed above indicate the arrival of an exciting era but also a crossroad for digital technology and language learning. Significant gaps remain both theoretically and empirically in the understanding of how digital technologies may be leveraged to enhance student performance, not just for language learning but for all domains of learning. Previously we discussed several important features/affordances of digital learning, including interactivity and autonomy/control, but, without a theoretical understanding of the roles of these affordances in learning, we will remain unclear about why and how DLL can benefit students and teachers. For example, what features in DLL environments are critical and conducive to L2 learning, and what empirical evidence is there? Does DLL learning lead to deeper cognitive processing and better L2 achievement than the traditional learning methods? Can DLL learning enable direct mappings between L2 and concepts and hence promote embodied representation in the L2? Are joint social attention and affective-emotional processing similarly important for adult L2 learning as for child L1 learning? What positive brain changes might we expect as a function of DLL, and what neural networks underlie DLL versus traditional L2 learning? And, finally, what emerging technologies in AI and big data analytics can we incorporate into DLL for personalized L2 learning? These are the kinds of questions that we as educators and researchers must tackle, and the answers may also have implications for better pedagogical practices and DLL product designs.

To address these questions, we must not only focus on the cognitive and social aspects of DLL, as already suggested by Peterson (2016; see 3.3). We must also study other dimensions of learning that may be critical for successful L2 learning. Below we discuss four such dimensions – namely, cognitive, social, affective, and neural – with respect to DLL.

4.1. Cognitive dimensions

An important area of study in cognitive science in the last decades has been embodied cognition. According to the embodied cognition theory (Barsalou, 2008; Glenberg, Sato, Cattaneo, Riggio, Palumbo & Buccino, 2008; Willems & Casasanto, 2011), our mental representations consist of not just symbolic abstractions, as assumed in classic cognitive theories, but conceptual properties that are deeply grounded in our body and our perceptions/actions in the physical world. Such theories highlight the "interaction between perception, action, the body and the environment" (Barsalou, 2008), and how body-specific (e.g., head, hands, feet) and modality-specific (e.g., auditory, visual, tactile) experiences are embedded in our mental representations. An embodied representation of a 'spoon' is not just its curved shape, the spelling of the letters, the fact it is used for eating, but an integrated memory of activity/eating with a spoon, the texture and size of a spoon, the fact it appears together with a plate or bowl, and that it is usually in a kitchen or restaurant, all of which form the conceptual representation of spoon – that is, the schema for 'spoon.' Furthermore, such embodied representations can activate the brain's visual and sensorimotor regions when the concept is retrieved, due to the way the concept has been encoded via perception and action.

The embodied cognition hypothesis allows us to see why DLL is fundamentally different from traditional classroom-based, translation-based, and teacher-centered L2 learning. In classroom-based vocabulary learning, for example, the teacher provides a list of foreign language words, and asks the student to learn by associating the list with the corresponding L1 word list, most likely through L2-to-L1 word translations; in traditional CALL, such translation-based associations can be implemented through digital flashcards, so that the correct associations can be tallied electronically. Learning in this way can be highly effective in the short term, but might result in the so-called 'parasitic' L2-on-L1 representation (Hernandez, Li & MacWhinney, 2005) or stronger L2-to-L1 links (Kroll & Stewart, 1994). This contrasts with the situation in which the child learns the L1 words; for example, the child acquires an embodied representation of 'spoon' through using the spoon in the kitchen, feeling its shape and texture, eating with it with a bowl, and often with a parent/adult around. Such perception-action features are absent in the classroom during adult L2 learning of the Spanish word 'la cuchara' through translation/association with its L1 equivalent 'spoon.' DLL can help to remedy this situation through technologies such as VR and simulated actions within VR that the learner can perform, as illustrated in Figure 1: the L2 learner can see, point to, pick up, and move kitchen objects associated with the L2 word/label, or even simulate the corresponding action (e.g., drinking with a cup, squatting to pick up a broom). Thus, DLL enables a child-like learning process, which may be critically

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press



Figure 1. Perception and action in immersive VR. (A) The L2 learner uses the handset to *point to* any item in the VR kitchen, which triggers the sound of the corresponding L2 word, in this example, 'dao' (knife in Chinese); (B) the learner virtually *picks up* and *moves* any object by pressing a trigger button with the index finger, in this example, a broom; (C) the learner *holds* a funnel to move it around; (D) the learner *opens* the refrigerator; (E) the learner uses a VR treadmill to navigate a virtual zoo; and (F) kangaroos in the virtual zoo, and as in (A) the learner uses the handset to point to the animal to trigger the L2 sound.

important for building an embodied representation in the L2: the learner encodes the L2 word by making direct contact with the concept without the mediation of L1, unlike in L2-to-L1 translation/association learning.

Relevant to the discussion here is the question of what type of perception and action will be most conducive to the establishment of embodied representations. According to the National Academies of Sciences, Engineering, and Medicine (2018), learning technologies offer 'affordances' (features or properties of objects that present a given object in a particular way when being used), and consideration of the affordances of a given technology is important for understanding student learning. Interactivity, adaptivity, feedback, linked representations, and communication with others are among the key affordances of today's digital technologies. Software designers as well as researchers should consider these affordances when developing or examining DLL products. For example, interactivity can be achieved in MALL, VR, and GBLL through user-to-user, user-to-object, or user-to-context interactions, and can be simulated with or without actual bodily actions (e.g., on the desktop computer, with a smartphone, or through *Microsoft Kinect;* see Lan et al., 2018). We will further analyze these affordances in the remainder of this article.

Learner characteristics are significant to our discussion of the cognitive dimensions, too. Identifying the cognitive abilities of the learner will enable us to understand how these abilities may be brought to bear on the L2 learning task. Specifically, two kinds of cognitive abilities have been implicated in technology-based learning: spatial abilities and executive function abilities (particularly working memory). Spatial abilities refer to an individual's ability to analyze spatial features of an environment, to navigate a complex landscape, and to construct a mental map. Various studies have shown that spatial abilities are essential for academic performance in a variety of science subjects (e.g., Kozhevnikov, Motes & Hegarty, 2007; Pani, Chariker & Naaz, 2013). For example, Naaz, Chariker and Pani (2014) found that students who scored higher on mental rotation tasks (Vandenberg & Kuse, 1978) also performed better on learning brain anatomy in a 3D dynamic environment. The abilities to mentally analyze and represent spatial features and relations are also highly relevant to language learning: the child learns L1 in a natural context with rich spatial cues, such as in environments of house, kitchen, and zoo, all involving a spatial layout with object locations relative to one another. DLL provides an authentic learning context for adult L2 comparable to that for child L1, aiming at grounding L2 learning in simulated or real environments (see Li & Jeong, 2020). Hsiao, Lan, Kao, and Li (2017) showed that, given the same DLL virtual spatial layout, L2 learners perform differently, both in the use of learning strategies (e.g., more self-exploratory roaming vs. sequential learning) and in the learning outcomes (high-

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

vs. low-achieving). Such differences may stem from learner characteristics, including spatial analytic abilities. Legault et al. (2019a) also showed that learners with higher spatial abilities performed better when learning in a VR zoo environment (where there is spatial navigation) than in a VR kitchen environment (where there is no spatial navigation). The authors further showed that, for highly successful learners, learning in VR vs. non-VR conditions did not matter, whereas, for the struggling learners, VR significantly promoted learning, a pattern consistent with data from GBLL-based research (see 3.3). Interestingly, such effects interacted with simulated action embodiment, such that, in general, kitchenware L2 names were learned better than animal names, perhaps due to the learner's ability to perform more action-based manipulations of objects in the virtual kitchen (where the learner can pick up and move objects around, which is not possible in the VR zoo). Figure 2 illustrates these differences based on Legault et al.'s (2019a) findings.

There has been ample evidence on the role of executive function, particularly working memory, in L2 learning (Baddeley, 2003; Miyake & Friedman, 1998; Wen, Biedrón & Skehan, 2017). However, it is so far unclear how working memory might play its role in DLL. Legault et al. (2019b) reported preliminary data regarding the neural correlates (see 4.4) of working memory for VR learning. These data suggested that working memory may be more important for DLL learners when the learning environment has many details and distractions (e.g., in a VR zoo), where the learner needs to attend to and monitor L2 target material while ignoring/inhibiting irrelevant information in the virtual environment. In such situations, the successful learner not only conducts more self-exploratory learning, but also dynamically keeps track of upcoming information using working memory and executive function.

Finally, previous work has indicated that DLL, as compared with traditional methods, can lead to deeper cognitive processing (e.g., Erhel & Jamet, 2013). In human memory research, the well-known 'encoding-specificity principle' (Tulving & Thomson, 1973) suggests that, if the encoding and retrieval contexts match, people learn better; for example, a word list encoded underwater would be retrieved better underwater than on dry land (Godden & Baddeley, 1975). Further, it is well established that deeper and more elaborative processing of that information (e.g., relating to semantic content of a word) leads to better long-term memory retention and retrieval, as compared to shallow processing (e.g., counting the number of letters in a word), supporting the classic cognitive theory of 'levels of processing' (Craik & Lockhart, 1972). Deeper processing may also involve multimodal processing, i.e., encoding of multiple sources of information (e.g., reading, writing, and hearing the same word). This 'multimodal advantage' is a central premise of the multimedia learning theory (Mayer, 2014), which suggests that students learn and remember better with words and pictures together than with words alone (see Liu, Wang, Li, Ding, Yang, & Li, 2020 for recent fMRI evidence). Multimedia platforms give the learner a chance to select, organize, and integrate diverse information, and mobile-based apps, VR, and game-based learning all take into consideration how auditory, visual, and tactile information may be leveraged simultaneously for successful L2 learning.

4.2. Social dimensions

DLL's role in promoting contextualized, situated, and embodied L2 representations is highlighted, by the above-discussed cognitive theories, cognitive abilities, and multimodal information processing. DLL, in essence, attempts to equate, with the help of technologies, conditions of adult L2 learning with those of child L1 learning by grounding the learning process in the context in which language is used. In learning the word 'spoon', the child abstracts a representation through repeated 'episodes' of interactions associated with using a spoon in the context; the same can be done through simulations in VR or games when adults acquire the corresponding L2 representation. The Unified Competition Model (MacWhinney, 2012) postulates that there are no fundamentally different principles underlying L1 learning vs. L2 learning, but the processes and contexts under which these two types of learning take place are different. If L1 and L2 learning conditions can be equated, L2 learners can fend off the 'risk factors' such as thinking in L1 (as opposed to using L2 for inner speech) and social isolation (as opposed to integrating socially and culturally with the L2 community). Recently, Caldwell-Harris and MacWhinney (2021) further expanded this view in an emergentist account of the age effect, focusing on how environmental support, cognitive abilities, and motivational factors change over time in children, adolescents, and adults.

The idea of action-based interactive learning is not new and has long been accepted in child language research (Meltzoff, Kuhl, Movellan & Sejnowski, 2009). For children, decontextualized situations (e.g., watching DVDs) do not induce learning; from the earliest stages infants already depend on social interaction, joint attention, shared intentionality, and eye-hand-body coordination for learning success (Kuhl, 2007; Tomasello, 2000; Yu & Smith, 2016). Researchers have realized that social interaction and joint attention may also be critical for L2 learning. Verga and Kotz (2017) showed that in simulated social learning in the lab, joint attention between the participant and the experimenter helps to orient the learner's attention to the correct meaning among competing alternatives. Caldwell-Harris, Goodwin, Chu and Dahlen (2014) compared adult L2 learning from live instructors versus that from videos and found that the physical/social presence of the teacher leads to better learning than when the teacher appears only in videos (consistent with findings from Kuhl, Tsao & Liu's 2003 infant study). These perspectives are highly consistent with both historical and recent trends in language acquisition and L2 learning, from sociocultural theory (Lantolf, 2006; Vygotsky, 1978) to usage-based language learning and processing (Tomasello, 2000, 2003), to input and interaction hypotheses (Krashen, 1988; Long, 1981), all of which highlight the properties and conditions in the learning environment, the linguistic input/output, and the interaction between these properties and learner-specific characteristics and cognitive profiles (see Ellis, 2019 and Mackey et al., 2012 for reviews and perspectives). A recent formulation of this interaction has been proposed by Claussenius-Kalman, Hernandez and Li (2021) in terms of the 3E framework, *Ecosystem*, *Expertise*, and *Emergentism*, which postulates that the emergent patterns of bilingual representation and cognitive processing reflect the dynamic interactions among the complex learning environment, the genotype of the individual, and the developing cognitive abilities of the learner.

On the basis of these data and theories, Li and Jeong (2020) proposed the 'social L2 learning' (SL2) hypothesis, according to which child L1-like representations can be achieved in L2 even for late adults through 'social learning' – learning that is perception and action-based, interactive, involving multimodal processing of information relevant to the target L2 environment, either through real-world or simulated contexts. One important SL2

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

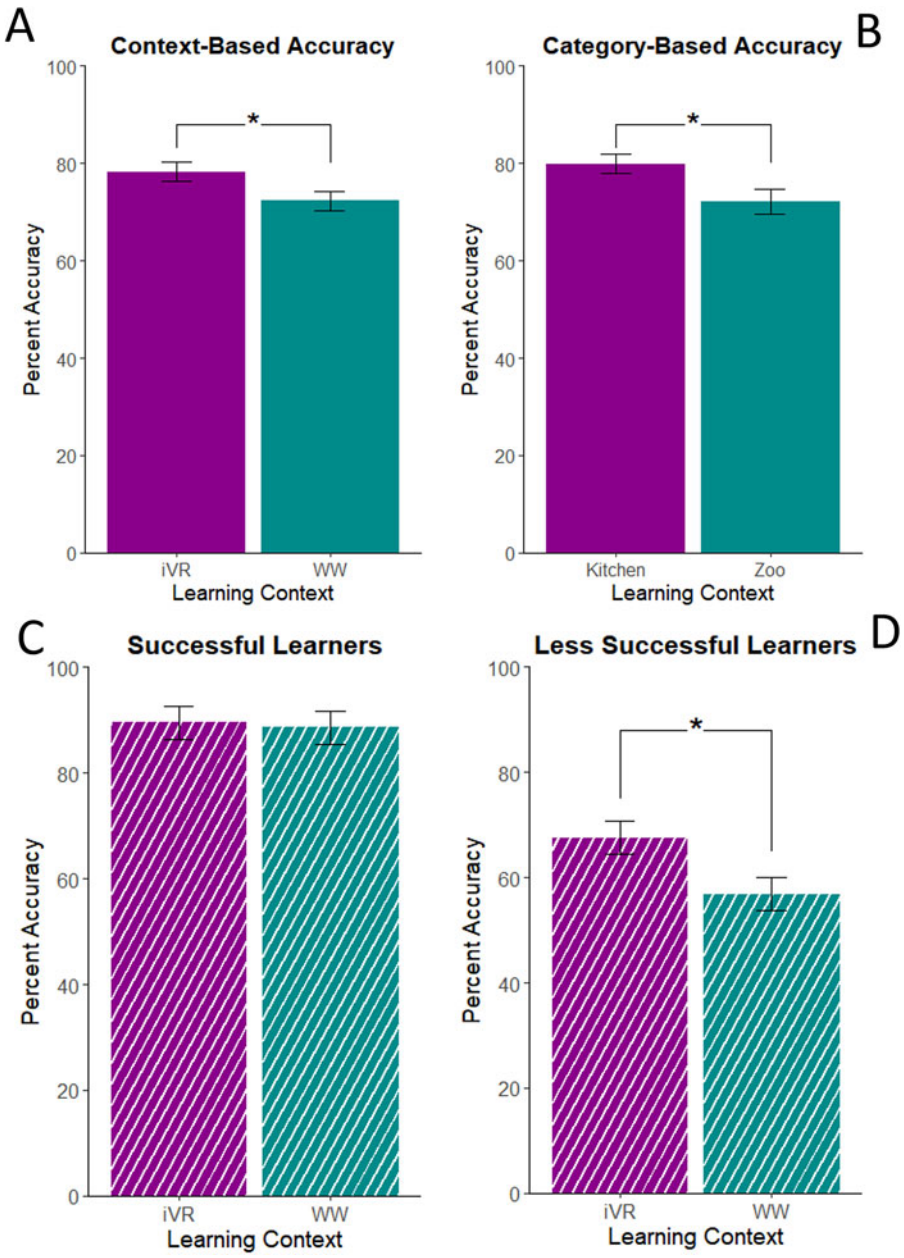


Figure 2. Effects of learning context, category, and individual differences. (A) There was an overall significant difference between immersive VR (iVR) vs. non-VR associative learning (WW, word-to-word association); (B) there was a significant difference between learning in Kitchen vs. learning in Zoo (both in iVR conditions); (C) there was no significant effect of learning context for Successful Learners; and (D) there was a significant effect of learning context for Less Successful Learners, with significantly higher accuracy in the iVR compared to the WW condition. Error bars indicate 95% confidence intervals and * indicates significant effect (based on Legault et al., 2019a).

hypothesis is that social learning can promote embodied L2 representations, because of the rich perceptual, sensorimotor, and affective-emotional processes that are embedded in the learning experience. Such experiences engage multimodal information integration, social reasoning, and motoric action or simulation, all of which reinforce long-term memory retention and facilitate retrieval. SL2 also provides a way for adult L2 learners to decouple the L2-to-L1 link that would otherwise be characteristic of late age of acquisition (the 'parasitic' representation; Hernandez et al., 2005; Li & Zhao, 2013). Moreover, such SL2 learning will necessarily recruit the brain's corresponding key regions that handle perception, action, and emotion, in both hemispheres (see 4.4). Given the social-affective as well as perception/action-based cues, social learning of L2 provides a genuine natural context comparable to that of L1 learning. Not surprisingly, the DLL platforms, most notably MALL, VR, and GBLL, all attempt to make the best use of such social cues for L2 learning. These cues may be analyzed with regard to 'affordances', important features that make the context be conducive to learning. Here we focus on two, interactivity and autonomy.

'Interactivity' in DLL means that the technology allows the learner to actively interact with the digital environment presented by the DLL platforms (e.g., with a virtual agent or avatar). For example, the learner can assume a specific role in a MMORPG gaming environment or have dialogues with a virtual agent in an immersive VR environment (e.g., *Mondly*™ relies on this method). Interactivity can also more broadly refer to any visual, manual, or bodily interactions with digital objects; for example, the learner can manipulate objects through hand movements (e.g., picking up a virtual cup in a kitchen) or bodily movements and locomotion (e.g., navigating a virtual town; see Figure 1E-F). Such interactivity is not social interaction in the strict sense but does engage perception/action-based learning in the context, in a way very different from reciting a list of word translations in

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

an L2 classroom. To the degree that a given digital technology enables interactivity, the technology offers different affordances and may consequently have different impacts on learning (e.g., desktop video games do not allow the user to conduct full-body movement during playing or learning, whereas immersive VR does).

In social learning, 'autonomy' (sometimes also called 'agency'; see Mayer, 2014) is another important affordance, implying that the learner is empowered to explore the learning environment, discover facts, control their own learning process and pace, and decide on what and how learning should proceed. This notion of learner autonomy has become particularly popular today, as the emphasis on student-centered learning has gradually taken center stage in education. In the L1 literature, there is evidence that even 9-month-old infants learn better when they have control of the presentation of speech materials for learning (Lytle, Garcia-Sierra & Kuhl, 2018). In traditional classrooms, the teacher provides the learning target and method; in flipped classrooms, the teacher serves as a facilitator and provides feedback; and in DLL learning, the learner decides on the learning goals (Egbert, Chao & Hanson-Smith, 2007), along with the order, time, and frequency with which the material will be acquired. The student will also have control of how he or she moves around in the digital environment (see the trajectory pattern analyses by Hsiao et al, 2017). The advantages conferred by autonomy in DLL are considerable, and the data derived from learner autonomy often provide information about learner characteristics, learning strategies, and L2 achievement outcomes that are otherwise unavailable (see Section 5).

4.3. Affective dimensions

As compared with research on the cognitive and social dimensions, relatively little work has been done to study the affective dimensions of DLL. However, it is clear from child L1 learning that affective processing, especially emotionality, is equally important for successful language learning. Lytle et al. (2018) argued that when children are learning with peers in the same environment, they show heightened social and emotional arousal, which motivates their learning and leads to better performance. Yu and Smith (2016) identified a positive correlation between child-parent joint attention to objects in the environment and the child's sustained attention, pointing to social interaction as the underlying factor that supports this correlation. It is important that social interactions involve a reciprocal affective relation: the child pays more attention to the object that the adult focuses on, the adult also provides a contingent response to the child's attention, which in turn increases the child's attention (i.e., sustained attention). Without such contingent responses and reciprocal interactions, there will be no role for social interaction to play in learning. Indeed, today's pandemic-induced online learning mode (e.g., through *Zoom* or *Microsoft Teams*) often lacks joint attention, contingent response, and reciprocal interaction between the students and the instructor. Sustained attention to the learning content is difficult to maintain in such a setting.

The SL2 hypothesis of Li and Jeong (2020) argues that lessons learned from child L1 are directly relevant to our understanding of adult learning of L2. As shown by Verga and Kotz (2017), even in L2, joint attention is important, but the underlying affective and emotional mechanisms, however, have not been fully explored. Our hypothesis is that social-affective cues could activate the learner's emotional responses as well as deeper cognitive processing, thereby facilitating learning and enhancing the quality of L2 representation. An important component of social learning is about how to better connect with others, both cognitively and emotionally, using joint attention and contingent responses. For example, eye contacts, facial cues, emotional expressions, hand and body gestures, are all human signals on top of textual and verbal information, serving as feedback, appraisal, and interests for continued engagement (or lack therefore); these are crucial to a regular face-to-face social interaction, as in child L1 learning, but are not usually available to classroom-based adult L2 learning. In particular, human faces serve a social function, carrying significant affective information: slight movements of our eyes, eyebrows, nose, lips, mouth, cheekbones, and chins can indicate subtle but important emotional states and convey meanings of happiness, anger, indifference, ignorance, or disgust. More recent studies have also shown that the perceived emotions from the instructor's face can serve as priming to the learner's positive or negative responses during learning (e.g., Lawson, Mayer, Adamo-Villani, Benes, Lei & Cheng, 2020; Pi, Chen, Zhu, Yang & Hu, 2020). The study of human facial expression has now become a burgeoning field in psychology and cognitive science (Calvo & Nummenmaa, 2016).

Given such significant affective functions of human faces, it is clear that under today's pandemic both the student and the instructor suffer when no reciprocal facial expressions are available in learning or teaching. It is also no surprise that the lack of affective processing in traditional L2 instruction may have led to the lack of affective representations of the acquired L2 material. In contrast to previous empirical emphases on how L2 learner's anxiety impedes learning, bilingual representation studies (see Dewaele, 2021, for a review) have shown that affective-specific feelings by emotion-laden words (words for affection, taboo words, swearwords, etc.) are more strongly evoked in L1 than in L2. This pattern could be due to the different contexts in which L1 vs. L2 is learned (in natural environments vs. in L2 classrooms) and the resulting semantic representation of emotions in L1 vs. L2 words. Importantly, such L1-vs.-L2 emotionality differences have been found most reliable when the L2 is a later-learned or less proficient/dominant language, showing that late adult L2 representations cannot easily incorporate the rich affective/emotional features that are typical of L1 representations (Caldwell-Harris, 2015). Pavlenko (2012) specifically linked L2 representation's weak emotionality to the decontextualized nature of traditional L2 classrooms where few opportunities are offered for integrating multimodal and multisensory information and where disembodied L2 representation results (see also 4.1).

DLL tools and platforms could potentially remedy the lack of L2 affective processing and emotionality differences through automatic feedback in MALL apps, avatars with emotional expressions in VR, and performance-contingent rewards in GBLL (Graesser et al., 2009; Park, Kim, Kim & Yi, 2019). Intelligent tutors or agents can also be built into DLL platforms using automatic speech recognition and AI, such that joint attention and contingent responses can be simulated (see D'Mello & Graesser, 2012 for incorporating human-like facial expressions in intelligent tutoring systems). However, simply providing the instructor's face images on a screen as in today's online teaching might not be sufficient: Resnik and Dewaele (2021) concluded in a recent study that the projection of the tiny 2D thumb-sized faces of teachers and peers on the screen does not convey the same emotional impact as do real human faces in student-teacher interactions. The Image Principle of the multimedia learning theory also states that "people do not necessarily learn more deeply

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

from a multimedia lesson when the speaker's image is added to the screen" (Mayer, 2014, p. 360).[3] Finally, whether real human faces and cartoonlike characters ('pedagogical agents'; see Section 5) make a difference to student learning is an active topic of investigation. Much work is needed in this area.

4.4. Neural dimensions

Our discussion has made it amply clear that DLL, due to its features/affordances on cognitive, social, and affective dimensions, enables L1-like representations in the L2, through the use of interactive and socially relevant contexts and multimodal/multisensory information. If there are such advantages of DLL, how does the brain reflect them? Despite much work in the study of the bilingual brain, we have so far very limited knowledge about how DLL tools and practices impact brain function and structure in L2 learning. Here we predict that the DLL methods will directly impact the L2 learning brain, and this prediction is based on converging evidence from two related literatures: a) action video game playing can enhance attentional control and cognitive resource allocation, leading to neuroplasticity in the central executive network (Bavelier, Green, Pouget & Schrater, 2012; Nahum & Bavelier, 2020); b) bilingual experience can increase executive function including attentional control, leading to brain changes also in the central executive network (Abutalebi & Green, 2007; Bialystok, Craik & Luk, 2012; Li, Legault & Litcofsky, 2014). There has also been recent neural evidence that game-based learning, as compared with non-game-based learning of the same material, leads to higher levels of activation in the brain's emotional and reward processing systems (Kober, Wood, Kiili, Moeller & Ninaus, 2020). Understanding the neural substrates of DLL will not only provide further evidence on the impacts of DLL, but also a window into how brain changes might result from the cognitive, social, and affective dimensions of DLL.

New evidence indicates that the brain can directly reflect the L1 vs. L2 difference with regard to embodied semantic representation: an integrated brain network that connects key language areas with semantic and sensorimotor regions is evoked when semantic processing is performed in L1, whereas such a network is absent or weakly configured for L2 processing (Zhang, Yang, Wang & Li, 2020). In the sensorimotor integration hypothesis of Hernandez and Li (2007), this difference results from the different ages of acquisition (AoA, early for L1 and late for L2). In the views of the declarative/procedural model of Ullman (2001), such difference is argued to be the result of procedural learning of L1 and declarative learning of L2. But according to the recent hypothesis of Li and Jeong (2020), such L1-L2 contrast is best seen as reflecting social learning for child L1 and association/translation learning for adult L2. There is already evidence in the literature that social learning in adult L2 can have a positive impact on the brain, measurable through functional and structural magnetic resonance imaging (MRI; see Stein, Winkler, Kaiser & Dierks, 2014 for an earlier review). For example, Jeong, Sugiura, Sassa, Wakusawa, Horie, Sato and Kawashima (2010) and Jeong, Li, Suzuki, Sugiura and Kawashima (2021) showed that words learned through videos of social interaction produced more activity in the right supramarginal gyrus (SMG) and angular gyrus (AG), whereas words learned through translation produced more activity in the left frontal gyrus (LFG). Verga and Kotz (2017) also showed that simulated partner interaction in L2 learning led to more brain activities in SMG and areas involved in visuospatial learning and sensorimotor processing.

However, there is so far little work focusing on the neural substrates of DLL in this direction. Hong et al. (2017) provided some preliminary evidence that child L2 English learners showed increased resting-state functional connectivity in Broca's and Wernicke's areas after a 12-week game-based training, but the study suffered from a small sample size and lack of a control group. A more recent study by Legault et al. (2019b) analyzed the structural MRI data from Lan et al. (2015), showing that L2 Chinese learners in the VR condition had a positive correlation between learning performance and brain structure in the right inferior parietal lobule (IPL), where brain structure was measured using cortical thickness. IPL has been regarded as a key hub for vocabulary learning and for multimodal information integration (Binder & Desai, 2011; Mechelli et al., 2004). By contrast, the learners in the non-VR condition (word-to-picture association) showed no such correlation.

Enabled by digital technology, DLL makes social-affective cues available to adult L2 learners that are normally only available to L1 learners. In other words, DLL enables social learning without putting the L2 learner in the physical social environment such as in immigration or study-abroad situations. The consequence is that DLL learners, as compared with translation/association-based learners, will necessarily engage a broader brain network in cortical, subcortical, and limbic systems, in both the left and right hemispheres, for effectively analyzing linguistic and non-linguistic perceptual information. This broadened brain network leads to enhanced cognitive processing, increased social-affective response, higher levels of motivation, better long-term memory retention, and faster memory retrieval. Figure 3 is an illustration of what such a network might look like.

This figure highlights the contribution of the right hemisphere to the learning of L2, contrasting the traditional left-hemisphere dominant language/lexical processes. It has become increasingly clear that the right hemisphere plays a much more important role than previously thought in adult L2 learning (see Qi & Legault, 2020, for a recent review). It is our hypothesis that DLL can enable the learner to establish direct and strong links between new L2 forms and social-affective features of the environment, leading to richly contextualized and embodied semantic representations. Much work needs to be done to identify such representations clearly in the L2 brain. We will need to rely on recent advances in network science (e.g., Bassett & Sporns, 2017) to delineate the specific connections, dynamic pathways, and overall organizations among the key brain regions, as well as the cooperation between the left and right hemispheres; in the case of DLL, we need to identify the particular impacts that MALL, VR, and GBLL may have on the structural brain change and functional connectivity due to L2 learning (see Li et al., 2014; Yang & Li, 2019; Zhang et al., 2020).

Given this perspective, future directions should also include the study of neural networks underlying social learning and their interactions with the extended language network (see Ferstl, Neumann, Bogler & Von Cramon, 2008; Hagoort, 2019; Meltzoff et al., 2009). For example, the learner may participate in a process of 'social reasoning', engaging the so-called 'theory of mind' (ToM; Frith & Frith, 2012; Saxe, 2006). ToM activates

[3]A sizeable literature exists in delineating the Image Principle by comparing the inclusion vs. non-inclusion of human faces in videos for multimedia learning (e.g., Atkinson, 2002; Craig, Gholson & Driscoll, 2002; Moreno, Mayer, Spires & Lester, 2001). The evidence remains mixed according to Mayer (2014).

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

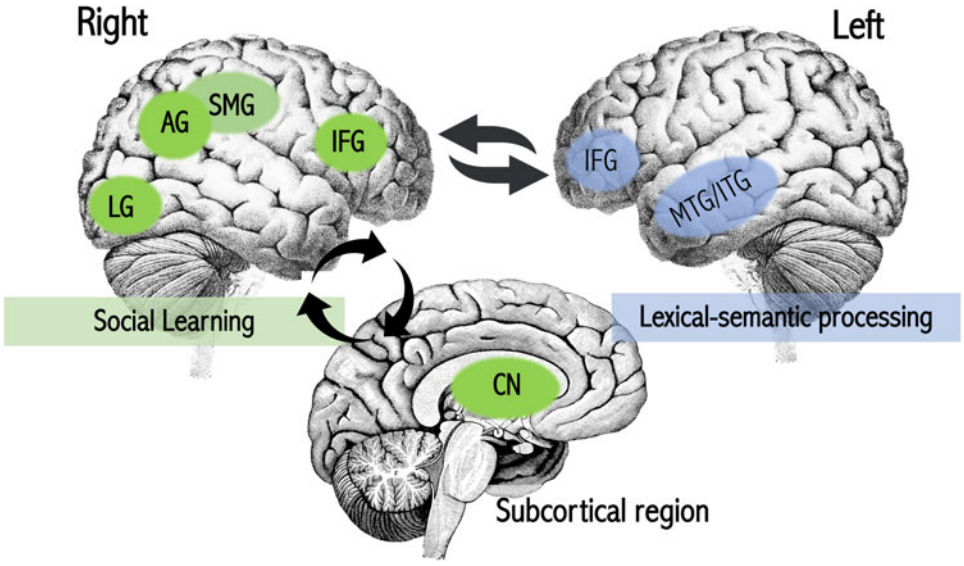


Figure 3. Brain network that supports lexical learning and social learning in both hemispheres. The figure illustrates a typical left-hemisphere lexical learning (blue) and a right-hemisphere social learning (green) system. The latter involves a right-heavy network that connects key regions in both hemispheres for visual processing (LG) and cognitive and linguistic processing (IFG, AG, SMG, MTG) with the subcortical region (CN for sequence learning). AG: angular gyrus; IFG: inferior frontal gyrus; SMG: supramarginal gyrus; LG: lingual gyrus; CN: caudate nucleus; MTG/ITG: middle/inferior temporal gyrus. (from Li & Jeong, 2020; with permission from *Springer Nature*)

the brain's mentalizing network, including medial prefrontal and bilateral temporoparietal junction regions, when thinking about other people's beliefs, desires, emotions, and intentions. In the case of language, this network may be engaged when the individual is trying to make inferences or take another person's perspective, which is highly relevant to the acquisition of L2 pragmatics that can also be aided by DLL (Sykes, 2017). Thus, we need to understand how our brain's linguistic system, memory system, emotional system, and theory of mind all work together as an integrated network to facilitate L2 learning and bilingual representation.

5. Emerging technologies and DLL: AI, Big Data, and personalized learning

The study of language learning has become a highly interdisciplinary enterprise due to its interaction with psychology, education, neuroscience, and now with machine learning. Meltzoff et al. (2009) used child language learning as a *bona fide* example to illustrate key principles for a 'New Science of Learning', in that language learning fulfills three premises simultaneously: (a) learning is a computational process, (b) learning is a socially interactive process, and (c) learning is supported by a dynamic neural circuitry linking perception and action. We believe that adult L2 learning can be equally positioned, if we adopt the DLL approach illustrated in this article. DLL follows the theoretical and methodological advances in education, cognitive science, and neuroscience, as discussed above. Moreover, DLL depends heavily on the latest technologies from mobile computing and VR to digital games. In this section, we discuss how emerging new technologies could further expand the impacts of DLL for the future.

Recent years have witnessed rapid developments and applications in AI and big data analytics. These developments have had profound impacts on all aspects of our lives. Although AI and data-driven language learning technologies are still at an early stage, learning with digital tools and platforms has become the norm as DLL attests, and it generates a vast amount of data in a short period of time (the so-called 'data deluge') which quickly exceeds the capacity of traditional data analytic methods. For example, in MALL, the apps can record each click as learning progresses; in VR, a student may traverse a virtual environment and every activity or movement may be recorded as a learning event (e.g., the activities depicted by Figure 1A-E); and in game-based learning, playing a game with multi-users could involve rapid interactive dialogues, resulting in many words and utterances in seconds. Further, cutting-edge immersive technologies such as VR-Eye integration and VR-EEG integration have enabled the collection of large-scale, multi-dimensional, and continuous data as learning occurs in real time, which include not only behavioral patterns but also eye gazes, electrophysiological, and neurocognitive responses during learning. Even learner's emotional and affective states/responses can be automatically captured through sensors and wearables (e.g., HTC Vive Facial Tracker, eye-trackers) or other experimentally designed tools (e.g., body posture measurement system, see D'Mello & Graesser, 2012). Such rich data provide, on a moment-by-moment basis, details about the object features that learners attend to, about learners' attention and cognitive spans, and about their spatial movements and navigation patterns in terms of time, speed, accuracy, and frequency. These complex multimodal and multimedia data differ significantly from traditional data collected after learning (answers to questionnaires and interviews, multiple choices, etc.), and lend themselves readily to data-intensive analytics based on advanced statistics, machine learning, and AI techniques.

One important question to ask is whether we can make use of the data deluge and data analytics from DLL to identify, predict, and adapt to individual differences in light of different learner characteristics. This is the idea of 'personalized learning' or 'precision education': educators take into consideration learner-specific characteristics, abilities, and strategies/styles of learning when developing curricula and pedagogies to fit the cognitive, social, and affective profiles and demands of different learners so as to

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

optimize learning (Hawk & Shah, 2007; see Luan et al., 2020 for a recent white paper on AI and big data in education). For example, corpus linguistic research has led to a large amount of word and text corpora, often open-access, with very detailed information about linguistic properties and usages, such as in the databases or corpora of WordNet (Fellbaum, 1998; Miller, 1995), BNC (BNC Consortium, 2007), and COCA (Davies, 2008). However, DLL tools and platforms have yet to seamlessly incorporate such information (e.g., lexical concordances) for intelligent L2 learning and teaching (see Ma, 2017 for a discussion).

To effectively design personalized learning, we need to understand both the internal characteristics of the learner (e.g., cognitive abilities, affective states, learning styles) and external characteristics of the environment (e.g., affordances of the learning context), and how the two interact (e.g., learning strategies in the context). In 4.1 we pointed out that individual differences in working memory may be particularly important for VR learning, but how working memory interacts with affordances of VR environments for L2 learning remains to be understood in the perspective of personalized learning. For example, Hsiao et al. (2017) showed how we could use advanced statistical analyses and computational models to identify the relations between navigation patterns of learners and their L2 learning strategies, and to predict their language learning success. In addition, using methods developed in other fields (e.g., 'roaming entropy' used to measure rat movements in maze running; Freund et al., 2013), we can also identify learners' traversing patterns within the digital environment. Such analyses indicated that the self-explorers ('high roamers') vs. the sequential learners ('low roamers') differed in learning outcome, high-achieving vs. low-achieving, respectively. Further, individuals with higher working memory, when facing a complex virtual environment, may be more able to keep track of the continuously updating visual scenes and ignore or inhibit irrelevant information, and therefore they are the ones more likely to adopt self-exploratory learning.

The next question to ask is if we might be able to modify and adapt the digital environment or virtual context to optimize individualized learning; for example, some distracting or 'seductive' details not directly relevant to the learning task can be simplified or eliminated in the virtual environment, such that individuals who have a lower working memory may more effectively focus on the L2 targets without getting distracted (see 4.1). This would make much sense in light of the 'cognitive load' theory (Mayer & Moreno, 2003; Sweller, 1994), according to which irrelevant audio-visual details (e.g., illustrations, images, faces), even if appearing highly attractive, can present increased demands on the learner's cognitive processing resources. However, we need to understand what audio-visual materials might be more distracting from learning versus more conducive to learning, and what kinds of learners might benefit more or less from them. As mentioned earlier, to design effective DLL tools and platforms, we must separate technological features from human characteristics and learner abilities, which will in turn help us better understand the efficacy of technological products. We need a greater synergy between technology and human characteristics – nowhere more than in education – and we must make our technologies be adaptive to individuals' cognitive, social, affective, and linguistic abilities and profiles.

How can we best combine the power of digital technology and that of AI and machine learning for developing personalized L2 education? Preliminary evidence suggests that we can indeed develop learner-specific models and materials through data-driven methods to enhance personalized vocabulary learning; for example, by analyzing detailed individual learning logs (e.g., Zou & Xie, 2018). One critical aspect, in addition to the key affordances of digital technologies discussed above, is feedback, which has been extensively examined in the multimedia learning literature generally (e.g., Moreno & Mayer, 2004; Moreno & Valdez, 2005) and in second language acquisition research specifically (Mackey et al., 2012; Presson et al., 2013). Feedback has been shown to contribute positively to learner motivation, cognitive processing, memory retention, and learner's enjoyment/feeling of rewards (e.g., Erhel & Jamet, 2013; Sweetser & Wyeth, 2005). In this respect, an exciting domain inspired by AI and big data analytics is the development of intelligent tutoring systems (ITS, such as AutoTutor; see Graesser et al., 2005; Nye, Graesser & Hu, 2014). ITS incorporates AI and machine learning algorithms to provide the learner with direct, immediate, and to-the-point feedback, not simply in the form of right or wrong answers. Like a human instructor, ITS can give feedback containing detailed, content-based corrections, comments, and suggestions, in response to and tailor-made to the individual's learning behavior and outcome.

Feedback represents a key affordance for digital technology to be both personal and humanistic – personal because it considers learner-specific patterns and humanistic because it incorporates other human-relevant features in the learning environment. In human face-to-face tutoring, the learner has social-affective-emotional cues including facial expressions, eye gazes, and body and manual gestures. ITS systems aim to incorporate, in addition to content-based feedback, such personal features through the design of animated 'pedagogical agents', the anthropomorphic animated human-like characters, to serve as virtual tutors. Johnson, Rickel and Lester (2000) and Johnson and Lester (2016) suggested that pedagogical agents should possess these social and affective-emotional features to qualify them as effective agents for guiding learning in interactive/immersive environments. Most important among these features, in our view, are the pedagogical agent's abilities to provide performance-contingent verbal and nonverbal feedback and to respond to affect and emotions in real time; hence, being socially intelligent (e.g., D'Mello & Graesser, 2012; Louwerse, Graesser, Lu & Mitchell, 2005).

Such features are particularly important for language learning (see also 4.3): without the ability to provide immediate feedback and affective responses, DLL tools will remain to be socially and emotionally distant to learners (and instructors). Unfortunately, existing 'intelligent language tutors' (ILTs) do not meet the standards yet (see Godwin-Jones, 2017 for a review), particularly given ILT's current focuses on providing corrective feedback on writing or giving text-based evaluations (see Shadiev & Yang, 2020). As an example, the popular VR software for L2 learning *Mondly*™ relies on a static stern-faced pedagogical agent responding to correct-vs.-wrong answers. Nevertheless, we see great potential in this domain given the significant advances in recent years in NLP (Hirschberg & Manning, 2015), automatic speech recognition (Golonka, Bowles, Frank, Richardson & Freynik, 2014; Li, Deng, Haeb-Umbach & Gong, 2015), affective computing (D'Mello & Graesser, 2010; Picard, 2015), and deep learning neural networks (LeCun, Bengio & Hinton, 2015). For example, automatic voice recognition can be built into the system to assess the learner's pronunciation accuracy and provide real-time feedback to the learner, which is already being explored by some commercial products (e.g., *Rosetta Stone*). We predict that AI-based tools will be further improved in the next few years, and be readily incorporated into or interfaced with MALL, VR, and GBLL to expand the utility and power of DLL.

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

In summary, there exist many opportunities and promises in leveraging AI and big data to make DLL more effective and personalized when we integrate the properties of the learning context including those from the environment, the tutor, and the learner. This integration will in turn facilitate the application of AI and big data analytics for better pedagogical design and language education. DLL represents an exciting interdisciplinary field where technology interfaces with human studies, and where theories and practices from cognitive science, neuroscience, and educational technology converge.

6. Conclusions

Language learning has entered a new era of pervasive digital applications. In light of the rapid developments in technology-enhanced education and AI-inspired innovations, DLL has become an exemplary interdisciplinary area of study and a gateway connecting language science, the society, and the industry. In this article, we have charted an overall picture of what DLL has evolved into, what impacts it has created, and what future promises it may hold. We have also attempted to provide theoretical perspectives from psychology, education, linguistics, and neuroscience to understand the cognitive, social, affective, and neural dimensions of DLL. DLL has enormous potential given the new generations of 'digital natives' and the interests in digital applications and blended learning in the foreseeable future. But significant work remains to be done to understand the mechanisms under which DLL might simulate language learning in its natural, authentic context and consequently enhance its learning success. There are also significant gaps that exist between our academic knowledge of student learning and the industry's commercial product design. We need quick knowledge transfer from academia to the industry, which is currently hindered by many factors, including bureaucracies at different levels, and such problems are exacerbated by the different paces adopted by the academia versus the industry. To mend such gaps, we need the academics to work more closely with the industry and with policy makers, which will facilitate and accelerate the development of both knowledge discovery and knowledge transfer (see Luan et al., 2020 for a discussion). We hope that integration of the emerging technologies with the science of learning will allow us to address not only the theoretical and practical problems associated with second language learning, but also unpredictable and long-term challenges posed by disruptive societal events such as the Covid-19 pandemic.

Acknowledgments. Preparation of this article has been partially supported by a grant from the Hong Kong Research Grants Council (Project # PolyU15601520) and a Research Startup Fund from the Hong Kong Polytechnic University to PL, and grants from the Ministry of Science and Technology of Taiwan (#MOST110-2511-H-003-038-MY3 and #MOST 109-2511-H-003-026-) to YJL. The authors wish to express gratitude to Chan-yuan Gu, Zhexiao Guo, Jennifer Legault, Yingying Peng, Jing Wang, Jiayan Zhao and other members in the Brain, Language, and Computation Lab for their assistance in the relevant projects, to Hyeonjeong Jeong and Brian MacWhinney for their helpful discussions, and to Sean McMinn for his comments on an earlier draft of the article.

References

Abutalebi, J, & Green, D (2007) Bilingual language production: The neurocognition of language representation and control. *Journal of Neurolinguistics* **20**(3), 242–275.

Acquah, EO, & Katz, HT (2020) Digital game-based L2 learning outcomes for primary through high-school students: A systematic literature review. *Computers & Education* **143**, 103667.

Al-Ahdal, A (2020) Using computer software as a tool of error analysis: Giving EFL teachers and learners a much-needed impetus. *International Journal of Innovation, Creativity and Change* **12**(2), 418–437.

Anderson, JR, Reder, LM, & Simon, HA (1996) Situated learning and education. *Educational Researcher* **25**(4), 5–11.

Atkinson, RK (2002) Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology* **94**(2), 416–427.

Baddeley, A (2003) Working memory and language: An overview. *Journal of Communication Disorders* **36**, 189–208.

Barsalou, LW (2008) Grounding symbolic operations in the brain's modal systems. In G. R Semin. & ER Smith (eds.), *Embodied Grounding: Social, Cognitive, Affective, and Neuroscientific Approaches*. Cambridge, UK: Cambridge University Press, pp. 9–42.

Bassett, DS, & Sporns, O (2017) Network neuroscience. *Nature Neuroscience* **20**(3), 353–364.

Bavelier, D, Green, CS, Pouget, A, & Schrater, P (2012) Brain plasticity through the life span: learning to learn and action video games. *Annual Review of Neuroscience* **35**(1), 391–416.

Bialystok, E, Craik, F, & Luk, G (2012) Bilingualism: Consequences for mind and brain. *Trends in Cognitive Sciences* **16**(4), 240–250.

Binder, JR, & Desai, RH (2011) The neurobiology of semantic memory. *Trends in Cognitive Science* **15**(11), 527–536.

BNC Consortium. (2007) *The British National Corpus* (version 3; BNC XML Edition, distributed by Bodleian Libraries, University of Oxford, on behalf of the BNC Consortium. URL: http://www.natcorp.ox.ac.uk/.

Caldwell-Harris, CL (2015) Emotionality differences between a native and foreign language: implications for everyday life. *Current Directions in Psychological Science* **24**, 214–219.

Caldwell-Harris, C, Goodwin, KS, Chu, E, & Dahlen, K (2014) Examining the advantage of a live instructor vs. video in a laboratory study. *Innovation in Language Learning and Teaching* **8**(3), 191–204.

Caldwell-Harris, CL, & MacWhinney, B (2021) Age effects in second language acquisition: Expanding the emergentist account. *Behavioral and Brain Sciences* (under revision).

Calvo, MG, & Nummenmaa, L (2016) Perceptual and affective mechanisms in facial expression recognition: An integrative review. *Cognition and Emotion* **30**(6), 1081–1106.

Casasanto, D, & Jasmin, KM (2018) Virtual reality. In AMB de Groot & P Hagoort (eds.), *Research methods in psycholinguistics and the neurobiology of language: A practical guide*. Hoboken, NJ: John Wiley & Sons, pp. 174–189.

Cerezo, L, Baralt, M, Suh, BR, & Leow, RP (2014) Does the medium really matter in L2 development? The validity of CALL research designs. *Computer Assisted Language Learning* **27**(4), 294–310.

Chapelle, CA, & Sauro, S (eds.) (2017) *Technology and second language teaching and learning*. Hoboken, NJ: John Wiley & Sons.

Chen, JC (2016) The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. *Computers & Education* **102**, 152–171.

Chen, MH, Tseng, W, & Hsiao, T (2018) The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. *British Journal of Educational Technology* **49**(1), 69–77.

Chen, H, Yang, C, & Lai, K (2020) Investigating college EFL learners' perceptions toward the use of Google Assistant for foreign language learning. *Interactive Learning Environments*. doi: 10.1080/10494820.2020.1833043

Chun, DM (2019) Current and future directions in TELL. *Educational Technology & Society* **22**(2), 14–25.

Claussenius-Kalman, H, Hernandez, A, & Li, P (2021) Expertise, Ecosystem, and Emergentism: Dynamic developmental bilingualism. *Brain and Language* (in press).

Craig, SD, Gholson, B, & Driscoll, DM (2002) Animated pedagogical agents in multimedia educational environments: Effects of agent properties, picture features, and redundancy. *Journal of Educational Psychology* **94**(2), 428–434.

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

Craik, FI, & Lockhart, RS (1972) Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* **11**, 671–684.
Davies, M (2008) *The Corpus of Contemporary American English (COCA)*. https://corpus.byu.edu/coca/
Dede, C (2009) Immersive interfaces for engagement and learning. *Science* **323** (5910), 66–69.
Dede, C, Jacobson, J, & Richards, J (2017) Introduction: Virtual, augmented, and mixed realities in education. In D Liu, C Dede, R Huang & J Richards (eds.), *Virtual reality, augmented reality, and mixed reality in education*. Hong Kong: Springer, pp. 1–18.
deHaan, J, Michael Reed, WM, & Kuwada, K (2010) The effect of interactivity with a music video game on second language vocabulary recall. *Language Learning and Technology* **14**(2), 74–94.
Dewaele, JM (2021) Research into multilingualism and emotions. In GL Schiewer, J Altarriba & BC Ng (eds.), *Language and Emotion: An International Handbook*. Berlin: Mouton De Gruyter.
D'Mello, SK, & Graesser, A (2010) Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction* **20**(2), 147–187.
D'Mello, SK, & Graesser, A (2012) AutoTutor and affective AutoTutor: Learning by talking with cognitively and emotionally intelligent computers that talk back. *ACM Transactions on Interactive Intelligent Systems* **2**(4), 1–39.
Egbert, J, Chao, C.-C., & Hanson-Smith, E (2007) Introduction: Foundations for teaching and learning. In J Egbert & E Hanson-Smith (eds.), *CALL environments: Research, practice, and critical issues* (2nd edition). Alexandria, VA: TESOL. pp. 1–14.
Ellis, NC (2019) Essentials of a theory of language cognition. *The Modern Language Journal* **103**, 39–60.
Erhel, S, & Jamet, E (2013) Digital game-based learning: Impact of instructions and feedback on motivation and learning effectiveness. *Computers & Education* **67**, 156–167.
Fellbaum, C (ed.) (1998) *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
Ferstl, EC, Neumann, J, Bogler, C, & Von Cramon, DY (2008) The extended language network: a meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping* **29**(5), 581–593.
Foomani, EM, & Hedayati, M (2016) A seamless learning design for mobile assisted language learning: An Iranian context. *English Language Teaching* **9** (5), 206–213.
Freund, J, Brandmaier, AM, Lewejohann, L, Kirste, I, Kritzler, M, Kruger, A, Sachser, N, Lindenberger, U, & Kempermann, G (2013) Emergence of individuality in genetically identical mice. *Science* **340**(6133), 756–759.
Frith, CD, & Frith, U (2012) Mechanisms of social cognition. *Annual Review of Psychology* **63**, 287–313.
Gardner, JR (1984) Computer-Assisted Learning and In-Service Teacher Training. *British Journal of Educational Technology* **15**(3), 175–182.
Gharehblagh, NM, & Nasri, N (2020) Developing EFL elementary learners' writing skills through mobile-assisted language learning (MALL). *Teaching English and Technology* **20**(1), 104–121.
Glenberg, AM, Sato, M, Cattaneo, L, Riggio, L, Palumbo, D, & Buccino, G (2008) Processing abstract language modulates motor system activity. *Quarterly Journal of Experimental Psychology* **61**(6), 905–919.
Godden, DR, & Baddeley, AD (1975) Context-dependent memory in two natural environments: On land and underwater. *British Journal of psychology* **66**(3), 325–331.
Godwin-Jones, R (2017) Authoring language-learning courseware. In CA Chapelle & S Sauro (eds.), *Technology and second language teaching and learning*. Hoboken, NJ: John Wiley & Sons, pp. 348–363.
Godwin-Jones, R (2019) In a world of smart technology, why learn another language? *Educational Technology & Society* **22** (2), 4–13.
Golonka, EM, Bowles, AR, Frank, VM, Richardson, DL, & Freynik, S (2014) Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer Assisted Language Learning* **27**(1), 70–105.
Graesser, A, Chipman, P, Haynes, B, & Olney, A (2005) AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education* **48**(4), 612–618.
Graesser, AC, Chipman, P, Leeming, F, & Biedenbach, S (2009) Deep learning and emotion in serious games. In U Ritterfeld, M Cody, & P Vorderer (eds.), *Serious games: Mechanisms and effects*. New York and London: Routledge, Taylor & Francis, pp. 81–100.
Grgurović, M, Chapelle, CA, & Shelley, MC (2013) A meta-analysis of effectiveness studies on computer technology-supported language learning. *ReCALL* **25**(2), 165–198.
Grosjean, F (2019) *A journey in languages and cultures: The life of a bicultural bilingual*. Oxford, UK: Oxford University Press.
Hagoort, P (2019) The neurobiology of language beyond single-word processing. *Science* **366**, 55–58.
Hannibal Jensen, S (2019) Language learning in the wild: A young user perspective. *Language Learning & Technology* **23**(1), 72–86.
Hawk, TF, & Shah, AJ (2007) Using learning style instruments to enhance student learning. *Decision Sciences Journal of Innovative Education* **5**(1), 1–19.
Hernandez, A, Li, P, & MacWhinney, B (2005) The emergence of competing modules in bilingualism. *Trends in Cognitive Science* **9**, 220–225.
Hernandez, AE, & Li, P (2007) Age of acquisition: its neural and computational mechanisms. *Psychological Bulletin* **133**(4), 638.
Hirschberg, J, & Manning, CD (2015) Advances in natural language processing. *Science* **349**(6245), 261–266.
Hong, JS, Han, DH, Kim, YI, Bae, SJ, Kim, SM, & Renshaw, P (2017) English language education on-line game and brain connectivity. *ReCALL* **29**(1), 3–21.
Hsiao, IYT, Lan, YJ, Kao, C.-L., & Li, P (2017) Visualization analytics for second language vocabulary learning in virtual worlds. *Educational Technology & Society* **20**(2), 161–175.
Hsu, CT, Clariana, R, Schloss, B, & Li, P (2019) Neurocognitive signatures of naturalistic reading of scientific texts: a fixation-related fMRI study. *Scientific Reports* **9**(1), 1–16.
Hung, HT, Yang, JC, Hwang, GJ, Chu, HC, & Wang, CC (2018) A scoping review of research on digital game-based language learning. *Computers & Education* **126**, 89–104.
Hung, HC, Young, SSC, & Lin, CP (2015) No student left behind: A collaborative and competitive game-based learning environment to reduce the achievement gap of EFL students in Taiwan. *Technology, Pedagogy and Education* **24**(1), 35–49.
Jeong, H, Sugiura, M, Sassa, Y, Wakusawa, K, Horie, K, Sato, S, & Kawashima, R (2010) Learning second language vocabulary: neural dissociation of situation-based learning and text-based learning. *Neuroimage* **50** (2), 802–809.
Jeong, H, Li, P, Suzuki, W, Sugiura, M, & Kawashima, R (2021) Neural mechanisms of language learning from social contexts. *Brain and Language* **212**, 104874.
Johnson, WL, & Lester, JC (2016) Face-to-Face interaction with pedagogical agents, Twenty years later. *International Journal of Artificial Intelligence in Education* **26**(1), 25–36.
Johnson, WL, Rickel, J, & Lester, JC (2000) Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education* **11**, 47–78.
Junaidi, J, Hamuddin, B, Julita, K, Rahman, F, & Derin, T (2020) Artificial intelligence in EFL context: Rising students' speaking performance with Lyra Virtual Assistance. *International Journal of Advanced Science and Technology Rehabilitation* **29**(5), 6735–6741.
Kearney, M, Schuck, S, Burden, K, & Aubusson, P (2012) Viewing mobile learning from a pedagogical perspective. *Research in Learning Technology* **20**, 14406. doi: 10.3402/rlt.v20i0/14406
Kober, SE, Wood, G, Kiili, K, Moeller, K, & Ninaus, M (2020) Game-based learning environments affect frontal brain activity. *PLOS ONE* **15**(11), e0242573. https://doi.org/10/gkgm72
Kokoç, M, Akçapınar, G, & Hasnine, MN (2021) Unfolding Students' Online Assignment Submission Behavioral Patterns using Temporal Learning Analytics. *Educational Technology & Society* **24**(1), 223–235.
Kozhevnikov, M, Motes, M, & Hegarty, M (2007) Spatial visualization in physics problem solving. *Cognitive Science* **31**, 549–579.
Kroll, JF, & Stewart, E (1994) Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* **33**, 149–174.
Krashen, SD (1988) *Second language acquisition and second language learning*. Englewood Cliffs, NJ: Prentice-Hall International.

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

Kuhl, PK (2007) Is speech learning "gated" by the social brain? *Developmental Science* **10**(1), 110–120.

Kuhl, P, Tsao, FM & Liu, HM (2003) Foreign-language experience in infancy: effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences* **100**, 9096–9101.

Lai, C, & Zheng, D (2018) Self-directed use of mobile devices for language learning beyond the classroom. *ReCALL* **30**, 299–318.

Lan, YJ (2014) Does Second Life improve Mandarin learning by overseas Chinese students? *Language Learning & Technology* **18**(2), 36–56.

Lan, YJ (2015) Contextual EFL learning in a 3D virtual environment. *Language Learning & Technology* **19**(2), 16–31.

Lan, YJ (2016) The essential design components of game design in 3D virtual worlds: From a language learning perspective. In M Spector, BB Lockee & MD Childress (eds.), *Learning, Design, and Technology. An International Compendium of Theory, Research, Practice, and Policy*. Switzerland: Springer International Publishing, pp. 1–18.

Lan, YJ (2020a). Immersion, interaction and experience-oriented learning: Bringing virtual reality into FL learning. *Language Learning & Technology* **24**(1), 1–15.

Lan, YJ (2020b) Immersion into virtual reality for language learning. *Psychology of Learning and Motivation* **72**, 1–26. (Volume 72: *Adult and Second Language Learning*, Eds., K.D. Federmeier & H.-W. Huang).

Lan, YJ, Fang, S, Legault, J, & Li, P (2015) Second language acquisition of Mandarin Chinese vocabulary: Context of learning effects. *Education Technology Research and Development* **63**, 671–690.

Lan, YJ, Hsiao, IYT, Fang, WC, & Chen, NS (2018) Real body versus 3D avatar: The effects of different embodied learning types on EFL listening comprehension. *Educational Technology Research and Development* **66**(3), 709–731.

Lan, YJ, & Lin, YT (2016) Mobile seamless technology enhanced CSL oral communication. *Educational Technology & Society* **19**(3), 335–350.

Lantolf, J (2006) Sociocultural theory and L2: State of the art. *Studies in Second Language Acquisition* **28**, 67–109.

Lawson, AP, Mayer, RE, Adamo-Villani, N, Benes, B, Lei, X, & Cheng, J (2020) Recognizing the emotional state of human and virtual instructors. *Computers in Human Behavior* 106554.

LeCun, Y, Bengio, Y, & Hinton, G (2015) Deep learning. *Nature* **521**(7553), 436–444.

Lee, S, Lo, Y, & Chin, T (2021) Practicing multiliteracies to enhance EFL learners' meaning making process and language development: a multimodal Problem-based approach. *Computer Assisted Language Learning* **34**, 66–91.

Legault, J, Zhao, J, Chi, Y-A., Chen, W, Klippel, A, & Li, P (2019a) Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages* **4**(1), 13.

Legault, J, Fang, S, Lan, Y, & Li, P (2019b) Structural brain changes as a function of second language vocabulary training: Effects of learning context. *Brain and Cognition* **134**, 90–102.

Levy, M, & Stockwell, G (2006) *CALL dimensions: Options and issues in Computer-Assisted Language Learning*. New York, NY: Routledge.

Li, J, Deng, L, Haeb-Umbach, R, & Gong, Y (2015) *Robust automatic speech recognition: A bridge to practical applications*. Waltham, MA: Academic Press.

Li, P, Legault, J, Klippel, A, & Zhao, J (2020) Virtual reality for student learning: Understanding individual differences. *Human Behaviour and Brain* **1** (1), 28–36.

Li, P, Legault, J, & Litcofsky, KA (2014) Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex* **58**, 301–324.

Li, P, & Jeong, H (2020) The social brain of language: Grounding second language learning in social interaction. *npj Science of Learning* 1–9. doi:10.1038/s41539-020-0068-7

Li, P, & Zhao, X (2013) Self-organizing map models of language acquisition. *Frontiers in Psychology* **4**: 828. doi: 10.3389/fpsyg.2013.00828

Lin, JJ, & Lin, H (2019) Mobile-assisted ESL/EFL vocabulary learning: A systematic review and meta-analysis. *Computer Assisted Language Learning* **32** (8), 878–919.

Lin, CC, Lin, V, Liu, GZ, Kou, X, Kulikova, A, & Lin, W (2020) Mobile-assisted reading development: A review from the activity theory perspective. *Computer Assisted Language Learning* **33**(8), 833–864.

Liu, D, Dede, C, Huang, R, & Richards, J (eds.) (2017) *Virtual reality, augmented reality, and mixed reality in education*. Hong Kong: Springer.

Liu, GZ, Chen, JY, & Hwang, GJ (2018) Mobile-based collaborative learning in the fitness center: A case study on the development of English listening comprehension with a context-aware application. *British Journal of Educational Technology* **49**(2), 305–320.

Liu, Z, Moon, J, Kim, B, & Dai, C (2020) Integrating adaptivity in educational games: A combined bibliometric analysis and meta-analysis review. *Education Technology Research and Development* **68**, 1931–1959.

Liu, C, Wang, R, Li, L, Ding, G, Yang, J, & Li, P (2020) Effects of encoding modes on memory of naturalistic events. *Journal of Neurolinguistics* **53**, 100863.

Loewen, S, Crowther, D, Isbell, D, Kim, K, Maloney, J, Miller, Z, & Rawal, H (2019) Mobile-assisted language learning: A Duolingo case study. *ReCALL* **31**(3), 293–311.

Lomicka, L, & Ducate, L (2021) Using technology, reflection, and noticing to promote intercultural learning during short-term study abroad. *Computer Assisted Language Learning* **34**(1–2), 35–65.

Long, MH (1981) Input, interaction, and second-language acquisition. *Annals of the New York Academy of Sciences* **379**, 259–278.

Louwerse, MM, Graesser, AC, Lu, S, & Mitchell, HH (2005) Social cues in animated conversational agents. *Applied Cognitive Psychology* **19**(6), 693–704.

Luan, H, Geczy, P, Lai, H, Gobert, J, Yang, SJH, Ogata, H, Baltes, J, Guerra, R, Li, P, & Tsai, CC (2020) Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology* **11**: 580820. https://doi.org/10/ghs3jz

Lytle, SR, Garcia-Sierra, A, & Kuhl, PK (2018) Two are better than one: Infant language learning from video improves in the presence of peers. *Proceedings of the National Academy of Sciences of the United States of America* **115**(40), 9859–9866.

Ma, Q (2017) Technologies for teaching and learning L2 vocabulary. In CA Chapelle & S Sauro (eds.), *Technology and second language teaching and learning*. Hoboken, NJ: John Wiley & Sons, pp. 45–61.

Mackey, A, Abbuhl, R, & Gass, S (2012) Interactionist approach. In S Gass & A Mackey (eds.), *The Routledge handbook of second language acquisition*. New York: Routledge, pp. 7–24.

MacWhinney, B (2012) The logic of the unified model. In S Gass & A Mackey (eds.), *The Routledge handbook of second language acquisition*. New York: Routledge, pp. 211–227.

Mayer, RE (2005) Cognitive theory of multimedia learning. In RE Mayer (ed.), *The Cambridge handbook of multimedia learning*. Cambridge, UK.: Cambridge University Press, pp. 31–48.

Mayer, RE (ed.) (2014) *The Cambridge handbook of multimedia learning* (2nd ed.). Cambridge, UK.: Cambridge University Press.

Mayer, RE (2016) What should be the role of computer games in education? *Policy Insights from the Behavioral and Brain Sciences* **3**(1), 20–26.

Mayer, RE, & Moreno, R (2003) Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist* **38**(1), 43–52.

Mechelli, A, Crinion, J, Noppeney, U, O'Doherty, J, Ashburner, J, Frackowiak, R, & Price, C (2004) Neurolinguistics: Structural plasticity in the bilingual brain. *Nature* **431**, 757.

Meltzoff, A, Kuhl, P, Movellan, J, & Sejnowski, T (2009) Foundations for a new science of learning. *Science* **325**, 284–288.

Miller, GA (1995) WordNet: A lexical database for English. *Communications of the ACM* **38** (11), 39–41.

Miyake, A, & Friedman, NP (1998) Individual differences in second language proficiency: Working memory as language aptitude. In A Healy & L Bourne (eds.), *Foreign language learning*. Mahwah, NJ: Lawrence Erlbaum, 339–364.

Mohsen, MA (2016) The use of computer-based simulation to aid comprehension and incidental vocabulary learning. *Journal of Educational Computing Research* **54**(6), 863–884.

Moreno, R, Mayer, RE, Spires, HA, & Lester, JC (2001) The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction* **19**(2), 177–213.

Moreno, R, & Mayer, RE (2004) Personalized messages that promote science learning in virtual environments. *Journal of Educational Psychology* **96**(1), 165.

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

Moreno, R, & Valdez, A (2005) Cognitive load and learning effects of having students organize pictures and words in multimedia environments: The role of student interactivity and feedback. *Educational Technology Research & Development* **53**(3), 35–45.

Naaz, F, Chariker, JH, & Pani, JR (2014) Computer-based learning: graphical integration of whole and sectional neuroanatomy improves long-term retention. *Cognition and Instruction* **32**(1), 44–64.

Nahum, M, & Bavelier, D (2020) Video games as rich environments to foster brain plasticity. In N Ramsey & J Millan (eds.), *Brain-computer interfaces* (Handbook of Neurology, Vol. 168), Elsevier, pp. 117–136.

National Academies of Sciences, Engineering, and Medicine (2018) *How people learn II: Learners, contexts, and cultures*. Washington, DC: The National Academies Press. doi:10.17226/24783.

Nicolaidou, I, Pissas, P, & Boglou, D (2021) Comparing immersive virtual reality to mobile applications in foreign language learning in higher education: a quasi-experiment. *Interactive Learning Environments*. doi:10.1080/10494820.2020.1870504

Nye, BD, Graesser, AC, & Hu, X (2014) AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education* **24**(4), 427–469.

Otto, S (2017) From past to present: A hundred years of technology for L2 learning. In CA Chapelle & S Sauro (eds.), *Technology and second language teaching and learning* (pp. 10–25). Hoboken, NJ: John Wiley & Sons.

Pandarova, I, Schmidt, T, Hartig, J, Boubekki, A, Jones, R, & Brefeld, U (2019) Predicting the difficulty of exercise items for dynamic difficulty adaptation in adaptive language tutoring. *International Journal of Artificial Intelligence in Education* **29**, 342–367.

Pani, JR, Chariker, JH, & Naaz, F (2013) Computer-based learning: Interleaving whole and sectional representation of neuroanatomy. *Anatomical Sciences Education* **6**(1), 11–18.

Park, J, Kim, S, Kim, A, & Yi, MY (2019) Learning to be better at the game: Performance vs. completion contingent reward for game-based learning. *Computers & Education* **139**, 1–15.

Pavlenko, A (2012) Affective processing in bilingual speakers: disembodied cognition? *International Journal of Psychology* **47**, 405–428 (2012).

Peeters, D (2019) Virtual reality: A game-changing method for the language sciences. *Psychonomic Bulletin & Review* **26**(3), 894–900.

Peterson, M (2010) Massively multiplayer online role-playing games as arenas for second language learning. *Computer Assisted Language Learning* **23**(5), 429–439.

Peterson, M (2016) The use of massively multiplayer online role-playing games in CALL: An analysis of research. *Computer Assisted Language Learning* **29**(7), 1181–1194.

Pi, Z, Chen, M, Zhu, F, Yang, J, & Hu, W (2020) Modulation of instructor's eye gaze by facial expression in video lectures. *Innovations in Education and Teaching International* 1–9.

Picard, RW (2015) The promise of affective computing. In RA Calvo, S D'Mello, JM Gratch & A Kappas (eds.), *The Oxford handbook of affective computing*. Oxford, UK: Oxford University Press, pp. 11–20.

Pikhart, M (2020) Intelligent information processing for language education: The use of artificial intelligence in language learning apps. *Procedia Computer Science* **176**, 1412–1419.

Presson, N, Davy, C, & MacWhinney, B (2013) Experimentalized CALL for adult second language learners. In JW Schwieter (ed.), *Innovative research and practices in second language acquisition and bilingualism*. Amsterdam: John Benjamins, pp. 139–164.

Puebla, C, Fievet, T, Tsopanidi, M, & Clahsen, H (2021) Digital language learning in older adults: Chances and challenges. *ReCALL* (in press).

Qi, ZH, & Legault, J (2020) Neural hemispheric organization in successful adult language learning: Is the left always right? *Psychology of Learning and Motivation* **72**, 119–163. (Volume 72: *Adult and Second Language Learning*, Eds., K.D. Federmeier & H.-W. Huang).

Rachels, JR, & Rockinson-Szapkiw, AJ (2018) The effects of a mobile gamification app on elementary students' Spanish achievement and self-efficacy. *Computer Assisted Language Learning* **31**(1–2), 72–89.

Reinders, H, & Wattana, S (2014) Can I say something? The effects of digital gameplay on willingness to communicate. *Language Learning & Technology* **18**, 101–123.

Reinhardt, J (2017) Digital gaming in L2 teaching and learning. In CA Chapelle & S Sauro (eds.), *Technology and second language teaching and learning*. Hoboken, NJ: John Wiley & Sons, pp. 202–216.

Resnik, P, & Dewaele, JM (2021) Learner emotions, autonomy and trait emotional intelligence in 'in-person' versus emergency remote English foreign language teaching in Europe. *Applied Linguistics Review* (in press).

Robertson, GG, Card, SK, & Mackinlay, JD (1993) Three views of virtual reality: Nonimmersive virtual reality. *Computer* **26**(2), 81.

Sadler, R (2017) The continuing evolution of virtual worlds for language learning. In CA Chapelle & S Sauro (eds.), *Technology and second language teaching and learning*. Hoboken, NJ: John Wiley & Sons, pp. 184–201.

Sandberg, J, Maris, M, & Hoogendoorn, P (2014) The added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning. *Computers & Education* **76**, 119–130.

Sato, T, Murase, F, & Burden, T (2015) Is mobile-assisted language learning really useful? An examination of recall automatization and learner autonomy. *Critical CALL – Proceedings of the 2015 EUROCALL Conference*, 495–501.

Saxe, R (2006) Uniquely human social cognition. *Current Opinion in Neurobiology* **16**(2), 235–239.

Shadiev, R, Hwang, WY, & Huang, YM (2017) Review of research on mobile language learning in authentic environments. *Computer Assisted Language Learning* **30**, 284–303.

Shadiev, R, Zhang, ZH, Wu, T.-T., & Huang, YM (2020) Review of studies on recognition technologies and their applications used to assist learning and instruction. *Educational Technology & Society* **23**(4), 59–74.

Shadiev, R, & Yang, M (2020) Review of studies on technology-enhanced language learning and teaching. *Sustainability* **12**(2), 524.

Shi, Z, Luo, G, & He, L (2017) Mobile-assisted language learning using WeChat instant messaging. *International Journal of Emerging Technologies in Learning* **12**(2), 16. https://doi.org/10/gkgfv3

Smith, GG, Li, M, Drobisz, J, Park, HR, Kim, D, & Smith, SD (2013) Play games or study? Computer games in eBooks to learn English vocabulary. *Computers & Education* **69**, 274–286.

Stein, M, Winkler, C, Kaiser, A, & Dierks, T (2014) Structural brain changes related to bilingualism: Does immersion make a difference? *Frontiers in Psychology*, **5**, 1116. https://doi.org/10.3389/fpsyg.2014.01116

Stockwell, G (2007) "Vocabulary on the move: Investigating an intelligent mobile phone-based vocabulary tutor." *Computer Assisted Language Learning* **20**, 365–383.

Suh, S, Kim, SW, & Kim, NJ (2010) Effectiveness of MMORPG-based instruction in elementary English education in Korea. *Journal of Computer Assisted Learning* **26**(5), 370–378.

Sundqvist, P, & Sylvén, LK (2012) World of VocCraft: Computer games and Swedish learners' L2 English vocabulary. In H Reinders (ed.), *Digital games in language learning and teaching*. London, UK: Palgrave Macmillan, pp. 189–208.

Sundqvist, P, & Wikström, P (2015) Out-of-school digital gameplay and in-school L2 English vocabulary outcomes. *System* **51**, 65–76. https://doi.org/10/f3n5jv

Sweetser, P, & Wyeth, P (2005) GameFlow: a model for evaluating player enjoyment in games. *Computers in Entertainment* **3**(3), 3–3.

Sweller, J (1994) Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction* **4**(4), 295–312.

Sykes, J (2017) Technologies for teaching and learning intercultural competence and interlanguage pragmatics. In CA Chapelle & S Sauro (eds.), *Technology and second language teaching and learning*. Hoboken, NJ: John Wiley & Sons, pp. 118–133.

Tai, TY, & Chen, HHJ (2020) The impact of Google Assistant on adolescent EFL learners' willingness to communicate. *Interactive Learning Environments*. doi:10.1080/10494820.2020.1841801

Tan, LH, & Xu, M (2020) Reading development in the digital age. *Human Behaviour and Brain* **1**, 71–73.

Tomasello, M (2000) The social-pragmatic theory of word learning. *Pragmatics* **10**, 401–413.

Tomasello, M (2003) *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tsai, YL, & Tsai, CC (2018) Digital game-based second-language vocabulary learning and conditions of research designs: A meta-analysis study. *Computers & Education* **125**, 345–357.

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press

Tu, Y, Zou, D, & Zhang, R (2020) A comprehensive framework for designing and evaluating vocabulary learning apps from multiple perspectives. *International Journal of Mobile Learning and Organisation* **14**, 370–397.

Tulving, E, & Thomson, DM (1973) Encoding specificity and retrieval processes in episodic memory. *Psychological Review* **80**(5), 352–373.

Ullman, MT (2001) The neuronal basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition* **4**(1), 105–122.

Van Deusen-Scholl, N (2015) Assessing outcomes in online foreign language education: What are key measures for success? *The Modern Language Journal* **99**(2), 398–400.

Vandenberg, SG, & Kuse, AR (1978) Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills* **47**(2), 599–604.

Verga, L, & Kotz, SA (2017) Help me if I can't: Social interaction effects in adult contextual word learning. *Cognition* **168**, 76–90.

Vesselinov, R, & Grego, J (2012) *Duolingo effectiveness study: Final report*. Queens College, City University of New York.

Vesselinov, R, & Grego, J (2016) *The Babbel efficacy study: Final report*. Queens College, City University of New York.

Vygotsky, L (1978) Interaction between learning and development. *Readings on the Development of Children* **23**(3), 34–41.

Wang, Y, & Christiansen, MS (2019) An investigation of Chinese older adults' self-directed English learning experience using mobile apps. *International Journal of Computer-Assisted Language Learning and Teaching* **9**(4), 51–71.

Wang, CP, Lan, YJ, Tseng, WT, Lin, YT, & Kao, CL (2020) On the effects of 3D virtual worlds in language learning- A meta-analysis. *Computer Assisted Language Learning* **33**(8), 891–915.

Ward, C, & Kennedy, A (1996) Crossing cultures: The relationship between psychological and socio-cultural dimensions of cross-cultural adjustment. In J Pandey, D Sinha, & DP Bhawuk (eds.), *Asian contributions to cross-cultural psychology*. Sage Publications, pp. 289–306.

Warschauer, M (2004) Technological change and the future of CALL. In S Fotos & C Brown (eds.), *New perspectives on CALL for second language classrooms*. Mahwah, NJ: Lawrence Erlbaum, pp. 15–25.

Wen, ZE, Biedrón, A, & Skehan, P (2017) Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching* **50**(1), 1–31.

Willems, RM, & Casasanto, D (2011) Flexibility in embodied language understanding. *Frontiers in Psychology* **2**, 116. doi:10.3389/fpsyg.2011.00116

Wimmer, J (2008) The multiple social meanings of digital games. What the first person shooter case study reveals us about the prerequisites for research. In N Carpentier *et al.* (eds.), *Democracy, journalism and technology: New developments in an enlarged Europe*. Tartu: Tartu University Press, pp. 335–42.

Yang, CY, Chen, I, & Ogata, H (2021) Toward precision education: Educational data mining and learning analytics for identifying students' learning patterns with ebook systems. *Educational Technology & Society* **24** (1), 152–163.

Yang, J, & Li, P (2019) Mechanisms for auditory perception: A neurocognitive study of second language learning of Mandarin Chinese. *Brain Sciences* **9** (6), 139.

Yang, SJH (2021) Guest editorial: Precision education - A new challenge for AI in education. *Educational Technology & Society* **24** (1), 105–108.

Yeh, YL, & Lan, YJ (2018) Fostering student autonomy in English learning through creations in a 3D virtual world. *Educational Technology Research & Development* **66**(3), 693–708.

Yu, C, & Smith, LB (2016) The social origins of sustained attention in one-year-old human infants. *Current Biology* **26**(9), 1235–1240.

Yu, Z (2018) Differences in serious game-aided and traditional English vocabulary acquisition. *Computers and Education* **127**, 214–232.

Zhang, X, Yang, J, Wang, R, & Li, P (2020) A neuroimaging study of semantic representation in first and second languages. *Language, Cognition and Neuroscience* **35**(10), 1223–1238.

Zou, D, Huang, Y, & Xie, H (2019) Digital game-based vocabulary learning: where are we and where are we going? *Computer Assisted Language Learning* **0**(0), 1–27.

Zou, D, & Xie, H (2018) Personalized word-learning based on technique feature analysis and learning analytics. *Educational Technology & Society* **21** (2), 233–244.

https://doi.org/10.1017/S1366728921000353 Published online by Cambridge University Press