Talking Head Animation as Pedagogical Agent in Language Learning: A Review on Instructional Strategy and Media

Kogilathah Segaran[1], Ahmad Zamzuri Mohamad Ali[2*] and Tan Wee Hoe[3]

[1,2,3]Faculty of Art, Computing and Creative Industry
Universiti Pendidikan Sultan Idris
Tanjong Malim, 35900, Perak, Malaysia
[*]zamzuri@fskik.upsi.edu.my

Abstract

Multimedia, particularly animation, has made significant contributions in educational settings, as it makes the learning process easier and more effective. Series of empirical studies performed over the last decade indicate that animation can provide various instructional roles, including improving the learning of English as a second language (ESL). Apparently, many ESL learners had difficulty in pronouncing certain English words, and natural English communication starts with proper pronunciation. These pronunciation issues in turn could lead to difficulty in finding a job because good communication skills is a principal criterion set by most if not all employers. In recent years, animated talking head technology has been developed to address these learning difficulties, which has set the path for the emergence of Computer-Assisted Language Learning (CALL), which applies a pedagogical agent or a virtual tutor in the form of an animated talking head. However, this methodology has rarely been researched in mobile-based educational settings, specifically related to animated talking heads. From what has been previously observed, mobile learning has recently become more popular in aiding distance education. Thus, this paper critically reviews the literature related to the use of animation and technology in aiding language learning, specifically among non-native English Speakers.

Keywords: animation, CALL, language learning, animated talking head, MALL

Introduction

It is generally known that multimedia, particularly animation, has played an important role in language learning. It has made significant contribution to the language learning process among various age groups of learners (Tamburini and Paci, 2002), particularly 3D animated talking head of virtual teachers in computer-assisted language learning applications (Wik, 2011; Wik and Hjalmarsson, 2009; Voce and Hamel, 2001). 3D animated talking head may be an essential instructional tool in supporting language learning through pronunciation modelling among non-native speakers (Badin et al., 2010).

This condition is true because non-native speakers face difficulty speaking English if they have poor pronunciation skills (Fraser, 2000). The difficulty arises among those who choose to study English after completing school (Gilakjani and Mohammad Reza, 2011). Much effort has been made in education to address this issue, including integrating language learning with multimedia and technology. Today, with rapidly evolving mobile technology, mobile learning or m-Learning offers a new approach in aiding education. Academia and the mobile world have set a new path for language learning, which has led to the introduction of Mobile-Assisted Language Learning (MALL) (Kukulska-Hulme and Shield, 2008). Since its inception, research and forum discussions on the topic of MALL have been conducted to identify effective ways of implementing and utilising this tool for successful language education.

Animation in Language Learning

Debate on animation typically relates to movies, cartoons or special effects. However, research has revealed that animation has significant contribution to education (Balasubramanyam, 2012; McMenemy and Ferguson, 2009; Doyle, 2001). Studies show that implementing animation in learning has led to positive outcomes for decades (Williamson and Abraham, 1995). Animation plays a role in improving the learning process, particularly in promoting thorough understanding of the subject matter (Mayer and Moreno, 2002). Today, animations are incorporated into computer-based multimedia learning and aid in many subject matters, including language learning (Cheng Lin and Fang Tseng, 2012; Kayaoğlu, Dağ Akbaş and Öztürk, 2011; Sundberg, 1998).

Language learning has undergone years of improvement, particularly after inserting technology-based methods in the eighties. This included the use of film, radio, television, language labs with audio/video tapes, computers, and interactive video (Cunningham, 1998). Traditionally, students studying English as a second language had to rely solely on text and audio materials such as cassette tapes, records and radio (Xiao and Jones, 1995). However, more recently, the use of film and video in learning English has provided a new method for acquiring it as a second language (Xiao and Jones, 1995). Unfortunately, this simultaneously led to lower engagement in student learning due to the lack of interactivity in the traditional way of acquiring English as a second language (Xiao and Jones, 1995). At this point in time, animation has evolved to aid in learning English as a second language as well as English as a foreign language.

This result has been proven in studies related to acquiring English as a second language. In a study by Lin, Chen and Dwyer (2006) on the effects of static visuals versus computer-generated animations, English language classrooms showed that computer-generated animated visuals, rather than static visuals, resulted in a more positive impact on immediate and delayed achievement. This was an appropriate strategy when the instructional objectives were to produce lower level learning outcomes involving factual knowledge, such as memorisation of vocabulary or application of simple grammar rules in composing sentences (Lin, Chen and Dwyer, 2006). Additionally, a study by Choi and Clark (2006) on cognitive and affective benefits of an animated pedagogical agent indicates that the usage of such tools provides entertainment and engagement in English language studies.

Analysing the animated pedagogical agent in promoting language learning has engendered an increased interest in educational research (e.g., Atkinson, 2002; Baylor and Ruy, 2003; Moreno and Mayer, 2000). Pedagogical agents are animated characters designed to function in educational settings to facilitate learning (Shaw, Johnson and Ganeshan, 1999). These agents are known as talking heads and model speech, facial expressions and gestures that support pedagogical strategies (Graesser, Chipman and King, 2008). 3D animated talking heads are now widely used for web services or as substitutes for face-to-face instruction (Lun, n.d.). Additionally, these talking heads appear as virtual tutors or teachers on learners' computer screens and contribute to various aspects of

language learning, including reading, pronunciation, conversation and practice (Busa, 2008). The use of 3D talking heads subsequently improved the learning process of acquiring a new language (Chen and Massaro, 2011).

Acquiring a new language is a significant challenge among all types of students and therefore, animated virtual tutors hold potential for utilising differentiated methods to support language studies (Massaro, 2006a). Many individuals have language disabilities and these students need additional instruction in language learning (Massaro, 2006a). However, these needs have not been met due to the lack of teachers and professionals equipped to give them individual attention (Massaro, 2006a). To address these barriers, students have found alternative ways to overcome this issue, such as using books or other alternative media. Unfortunately, these alternatives are not customised to a particular individual's needs, which is where the animated language tutor or animated talking head was introduced (Massaro, 2006a). As well, the face is an essential part of conveying a message through body language (Massaro, 2006b). Visual information delivered through the movements of the lips, tongue and jaws enhances audio comprehension in a noisy environment (Massaro, 2006b). Thus, visual information may play an important role in helping learners to distinguish words otherwise difficult to achieve with audio alone (Jesse and Massaro, 2010).

Speech is supported by facial expressions, emotions and gestures produced by a speaker (Massaro, 1998). In this regard, an experiment was conducted by Liu et al. (2007) on utilising visual speech for Chinese pronunciation training. 101 students in an introductory Chinese course at Carnegie Mellon University participated in the experiment. The students used a web-based learning environment. Students observed using a 3D animated talking head to learn final pronunciations rather than audio alone achieved significant improvement. Through the study, Liu et al. (2007) concluded that visual speech provides learners significant support in improving their pronunciation. However, one might question the use of text as verbal support in a talking-head application. Excluding text in pronunciation learning might cause difficulty among learners in identifying syllable breaks for proper pronunciation (Ahmad Zamzuri and Kogilathah, 2013). Thus, identifying solutions for verbal support in talking-head applications are necessary.

Talking Head Animation and Verbal Support

Animations have been integrated in educational technologies at a growing rate since the early 1980s (Ainsworth, 2008). There are several reasons for developers and researchers to use animations in their instructional design. One of the reasons might include the level of cognitive demand needed for a learning task (Tversky, Morrison and Bétrancourt, 2002). In many of Mayer's studies on the connection between animation and cognition, it has been proven that students learn more deeply from animation and narration compared to narration alone (Ainsworth, 2008). According to Schnotz and Rasch (2005), there are two ways that animations might support cognitive processing. The first way is to enable the function of animation, which occurs when animations provide additional information that cannot be displayed in pictures. The second method is in the facilitating function, when animations are able to help learners build mental models of situations with external support. This shows that animations make cognitive processing easier.

In addition to these, the learner engages in three important cognitive processes (Mayer and Moreno, 1998). The first cognitive process is in selecting where to apply incoming verbal information to support to text based input and where to apply incoming visual information to support image based input (Mayer and Moreno, 1998). The second cognitive process is in organising where to apply the word base to create a speech-based model of the system and where to apply the image base to create a visually based model of the system (Mayer and Moreno, 1998). Finally, the third process is in integrating, which occurs when the learner builds connections between related events in the speech-based model and the visually based model (Mayer and Moreno, 1998). Consequently, this relates to Mayer's modality principle. To test the modality influence, Moreno and Mayer (1999) conducted two experiments involving an animation depicting the process of lightning. In experiment 1, students viewed on-screen text presented near the animation or far from the animation, or simultaneously listened to a narration (Moreno and Mayer, 1999).

In experiment 2, they viewed on-screen text or listened to a narration, viewed on-screen text following or preceding the animation, or listened to a narration following or preceding the animation (Moreno and Mayer,

1999). Both experiments revealed a modality effect in which students learned more efficiently when verbal input was presented through speech rather than visually as text (Moreno and Mayer, 1999). However, it was recently shown that under certain conditions, visual texts can be preferable (Stiller et al., 2009). Instructional pacing seems to be one of the conditions that demonstrated this effect (Stiller et al., 2009). Stiller et al. (2009), investigated the effects of pacing and text modality on cognitive load and performance. The study proved that visual text instruction was the most efficient (Stiller et al., 2009). This condition can also be applied to the 3D animated talking head with audio and text applications. To pronounce accurately, a learner should know the word being stressed and how it is shown in a syllable break. To pronounce the stress pattern of a word clearly, the correct number of syllables needs to be produced (Pronunciation, 2011). Therefore, to show this, the syllable break needs to be placed as text in the application. However, specific research is needed to further confirm this claim.

Facial Expression and Lip Syncing

Adding to the use of verbal support in the audio and text of a talking head, facial expression is necessary to make language learning more efficient and effective (Wik and Hjalmarsson, 2009). Previous studies in the field of neuroscience, cognitive science, and psychology specify that emotions play a significant role in attention, planning, reasoning, learning, memory, and decision making (Picard, 1997). Emotions also influence perception, cognition, coping, and creativity (Johnson, Rickel and Lester, 2000; Picard, 1997). In the teaching of pronunciation, non-verbal communication such as facial expressions and gestures has become essential (Brown, 2007). However, non-verbal tools, including tone of voice, body posture, facial expression, and gestures can help a teacher make the classroom interesting and motivate students to actively engage in classroom activities (Sime, 2006). These tools also enhance the students' learning capability and their ability to recall information (Allen, 2000; Lazaraton, 2004). This eventually allows students to retain pronunciation knowledge.

In fact, recent studies on learning pronunciation have introduced new methods and such as face-to-face communication (Rodgers, 2001). Face-to-face communication is one of 10 scenarios introduced by Rodgers

(2001), which engages all aspects of human communication such as facial expression, gesture, tone and so forth to support teaching a second language in the 21st century (Rodgers, 2001). Additionally, lip synchronisation is one of the primary features of a talking head (Lun, n.d.). Because English is a language that depends upon airflow, lip shape, tongue position, teeth position and jaw movement (Baxter, 1993), the process of learning pronunciation can be practised by observing lip syncing activities (Sumby and Pollack, 1954; Benoıˆt and Le Goff, 1998).

Computer-assisted Language Learning

In addition to these strategies, instructional media plays a role in increased student learning. Technology-based training has increased since the introduction of computers 50 years ago (Graesser, Chipman and King, 2008). The extant research supports that students learn better in technology-rich learning environments than with classroom lectures, reading textbooks, and non-interactive control circumstances (Graesser, Chipman and King, 2008). Likewise, Malik and Shabbir (2008) and Saba (2009) emphasised the effective use of technology in creating new opportunities for independent learning as one method by which to increase student achievement. This situation also affected the way students acquire a second language. The use of these new technologies and multimedia elements for language teaching and learning has become a method of acquiring a second language, known as Computer-Assisted Language Learning (CALL) (Gamper and Knapp, 2002).

Advances in the Information and Communication Technology (ICT) have changed the roles of language teachers and learners (Park and Son, 2009) whereby it requires learners to be active participants in the learning process rather than passive recipients through their ability to control their own learning in a technology-enhanced learning environment (Brown, 1991). In a technology-enhanced environment, learners can manage their own learning process by collecting information and synthesising it independently (Park and Son, 2009). The classroom becomes more learner-centred when learners are able to make their decisions and work independently on the subject matter (Park and Son, 2009). Dunkel (1990) asserts that using computer technology as a tool can increase the language learner's self-esteem, vocational readiness, language proficiency and overall academic skills. In fact, many literature reviews corroborate the

benefits of CALL in acquiring second language. One review includes the exploration of the application of certain technologies in specific language areas (Liu et al., 2002). Hypermedia technologies such as linking and interactive capabilities have been discussed as a tool to enhance vocabulary learning (Liu, 1994) and reading comprehension (Hult et al., 1990). Moreover, use of video and audio are considered to be an advantage to support comprehension practice (Chun and Plass, 1997). On the other hand, multimedia technology represented in CALL is equipped to provide the enriching contexts that are important in language learning (Kramsch and Andersen, 1999).

Animation in CALL

Upon analysing multimedia technology in helping CALL, animation has played an important role. The trend of CALL and Computer-Assisted-Pronunciation-Training (CAPT) in supporting second language acquisition uses a virtual conversational agent or an animated talking head. This provides the ability to create a human-like tutor and is more beneficial than the desktop metaphor as an instructional interface (Wik and Hjalmarsson, 2009). Because CALL can be implemented outside of the classroom, animations and audio used in a CALL system has proven to be important elements of effective English language learning tools; it is suitable to be combined with standard curriculums and other appropriate design features to achieve what is naturally the outcome of a traditional classroom (Alqudah and Abdulsalam, 2013).

On the other hand, users interacting with animated tutors spend more time with the system because they believe it performs better and find the system interesting to interact with versus interacting with animated agents on a desktop interface (Walker et al., 1994; Koda and Maes, 1996; Lester and Stone, 1997; van Mulken and Andre, 1998). Moreover, speech is multimodal and communication is normally performed verbally as well as through facial expression (Wik and Hjalmarsson, 2009). This supports the idea that visual information supports speech perception (Sumby and Pollack, 1954). Virtual conversational agents or animated talking heads are able to give feedback on articulations that a human tutor finds it difficult to express or demonstrate (Wik and Hjalmarsson, 2009). Virtual reality displays of the face shows the position and movement of intra-oral articulators in combination with speech signals may improve the learner's

acquisition and production of new language sounds by analysing the relationships between speech sounds and gestures (Engwall, 2008).

Thus, CALL is moving slowly towards Mobile-Assisted Language Learning because mobile learning is the fast-growing element in current educational technology.

Mobile-assisted Language Learning

Mobile learning has grown rapidly in supporting an informal learning environment, which increases learning time outside the classroom (Derakhshan and Khodabakhshzadeh, 2011). Early generations of mobile learning projects preferred to offer formal activities carefully designed by educators and technologists as emerging technologies were not yet widely available or easy to understand (Pettit and Kukulska-Hulme, 2007). However, many of today's learners own mobile devices which enable them to easily engage in activities motivated by their personal needs and conditions of use, including activities needing to be completed while traveling and on the move (Kukulska-Hulme, Traxler and Pettit, 2007; Pettit and Kukulska-Hulme, 2007).

Today, mobile learning has a large impact on the secondary and tertiary learning processes by which many researchers have explored and studied the impact of techniques and the devices used. Mobile learning has also evolved as a medium of delivery from laptop to netbook and now to smart phones such as the iPhone and Android phones. In a paper presented by Herrington et al. (2010), researchers have concluded that introducing the iPhone and iPod Touch to the School of Education at Murdoch University became the medium of change to its traditional teaching and research. Schools in South Asia are also beginning to utilise this new technology to improve the standard and method of learning observed at Cempaka School, the first school in South Asia to receive the Apple Distinguished School status for innovative use of technology in the classroom in 2002 (Education on the Go, 2009). Rapid advancements in research and practice using mobile learning development shows that mobile technologies can be effectively used as communication tools when used by a wide range of learners in a mixture of settings (Kukulska-Hulme, 2010).

Mobile learning has given students the luxury of learning anytime and anywhere for any type of learner (Godwin-Jones, 2005; Kadyte, 2004; Kukulska-Hulme, 2005). At the same time, mobile learning is also one means of facilitating lifelong learning (Valk, Rashid and Elder, 2010). Mobiles can support a great amount of learning that occurs in between everyday activities, learning that occurs spontaneously in unprepared settings outside the classroom and outside of the usual environments of home and the office (Valk, Rashid and Elder, 2010). Mobile technologies enable learning that occurs across time and place as learners apply what they learn in one environment to developments in another (Sharples, Taylor and Vovoula, 2005; 2007). Mobile learning creates thus a situation in which education becomes learner centred, giving the learner greater responsibility for the learning process instead of passively receiving information from an instructor.

Colpaert (2004: 262) observed that throughout the history of CALL, periods of professional development have caused periods of beginner's application development and he wondered whether "the mobile hype will burst out as soon as tools become available allowing teachers and researchers to develop their own mobile applications and tools". This hype that Colpaert (2004) referred to is the rise of Mobile-Assisted Language Learning (MALL), which is rapidly becoming a reality as these tools are now owned by teachers and learners (Motteram and Stanley, 2011). With mobile phones linked to education, there exists a potential to provide a rich learning environment for the learners (Stockwell, 2010). In recent years, MALL appears to be a new teaching tool in the educational field (Chinnery, 2006). This is because researchers believe that in-class only activities are not sufficient for effective language learning and learners should be given opportunities to learn the language beyond classroom activities (Saran, Seferoglu and Cagiltay, 2009). Development of mobile technology has paved the path for educational improvement, particularly in mobile language learning. Research has been conducted to prove that mobile learning can enhance and ease the learning process for different groups of learners, especially in MALL (Idrus, 2011; Sood, 2010; Fotouhi-Ghazvini, Earnshaw and Haji-Esmaeili, 2009). The development of mobile and wireless technologies has broadened possibilities in the domain of language teaching (Joseph and Uther, 2009). Several methods have been used to improve the language learning using mobile

applications such as using multimedia messages in improving pronunciation (Saran, Seferoglu and Cagiltay, 2009).

Animation in MALL

Multimedia has a significant contribution in MALL. In recent studies, researchers have explored the use of multimedia messages via mobile phones for improving language learners' pronunciation of words (Saran, Seferoglu and Cagiltay, 2009). In a study conducted by Saran, Seferoglu and Cagiltay (2009), they concluded that when comparing three different modes of delivery including web based, hand outs and multimedia messaging service (MMS), students in the MMS delivery mode groups performed better than the web and hand out groups. The elements used in the MMS delivery mode were 2D animated images. However, animation, particularly 3D animations assisting language learning, was primarily a computer-based model such as the 3D talking-head system. There was research conducted on implementing 3D talking heads on mobile phones such as voice interactive services but there is little research completed on the use of 3D talking heads in mobile applications for assisting language learning. Therefore, continued research into 3D talking head mobile applications used in supporting language learning is essential.

Conclusion

Animation has made a significant contribution in the education industry among various content areas for the past several decades. It also plays an important role in the learning of a second language, determined through the conclusions of recent studies. Animation, specifically the 3D talking head, has been observed to be as effective a tool as a virtual teacher in aiding second language learning. This can be seen through the emergence of CALL and MALL, which include multimedia elements such as animated talking head technology in their systems or applications. However, little educational research has been completed using the mobile-based application concerning the 3D talking head. Thus, mobile learning is becoming more widespread in supporting current educational practices. Furthermore, MALL shows increased development in assisting language learning. Because of this, researching the effects of 3D talking head mobile applications in language learning seems essential.

Acknowledgement

References

Ahmad Zamzuri Mohamad Ali and S. Kogilathah. 2013. 3D Talking-head mobile app: A conceptual framework for English pronunciation learning among non-native speakers. *English Language Teaching* 6(8): 66–76.

Ainsworth, S. 2008. How do animations influence learning? In *Current perspectives on cognition, learning and instruction*: *Recent innovations in educational technology that facilitate student learning*, eds. D. Robinson and G. Schraw, 37–67. Information Age Publishing.

Allen, L. Q. 2000. Nonverbal accommodations in foreign language teacher talk. *Applied Language Learning* 11: 155–176.

Alqudah, M. and A. Abdulsalam. 2013. Basic English language tools for beginners: Using animations and audio. *International Journal of Scientific & Engineering Research* 4(4): 228–233.

Atkinson, R. K. 2002. Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology* 94: 416–427.

Badin, P., Y. Tarabalka, F. Elisei and G. Bailly. 2010. Can you 'read' tongue movements? Evaluation of the contribution of tongue display to speech understanding. *Speech communication* 52: 493–503.

Balasubramanyam, V. 2012. Animations in medical education. *Medical Journal of Dr. D.Y. Patil University* 5(1): 22. http://www.mjdrdypu.org/article.asp?issn=0975-2870;year=2012;volume=5;issue=1;spage=22;epage=22;aulast=Balasubramanyam (accessed 22 February 2013).

Baxter, B. 1993. *Studying tips: Speaking tips*. http://www.musicalenglishlessons.org/tips-speaking.htm (accessed 20 December 2011).

Baylor, A. and J. Ryu. 2003. Does the presence of image and animation enhance pedagogical agent persona? *Journal of Educational Computing Research* 28(4): 373–395.

Benoı̂t, C. and B. Le Goff. 1998. Audio–visual speech synthesis from French text: Eight years of models, designs and evaluation at the ICP. *Speech Communication* 26: 117–129.

Brown, A. 2007. The use of nonverbal features in teaching phonetics. Proceedings of the Phonetics Teaching & Learning Conference, UK.

Brown, H. D. 1991. TESOL at twenty-five: What are the issues? *TESOL Quarterly* 25: 245–260.

Busa, M. G. 2008. New perspectives in teaching pronunciation. In *Didactas to ecolingua: An ongoing research project on translation and corpus linguistics,* ed. A. Baldry, 165–182. Trieste: Edizioni Università di Trieste.

Cheng Lin, C. and Y. Fang Tseng. 2012. Videos and animations for vocabulary learning: A study on difficult words. *Turkish Online Journal of Educational Technology* 11(4): 346–355.

Chen, T. H. and D. W. Massaro. 2011. Evaluation of synthetic and natural Mandarin visual speech: Initial consonants, single vowels, and syllables. *Speech communication* 53: 955–972.

Chinnery, G. M. 2006. Emerging technologies going to the mall: Mobile assisted language learning. *Language Learning & Technology* 10(1): 9–16.

Choi, S. and R. E. Clark. 2006. Cognitive and affective benefits of animated pedagogical agents in learning English as a second language. *Journal of Educational Computing Research* 33(2): 455–480.

Chun, D. and J. Plass. 1997. Research on text comprehension in multimedia environments. *Language learning and technology* 1(1): 60–81.

Colpaert, J. 2004. From courseware to coursewear? *Computer Assisted Language Learning* 17(3–4): 261–266.

Cunningham, D. 1998. 25 years of technology in language teaching: A personal experience. Babel: *Journal of the Australian Federation of Modern Language Teachers' Associations* 33(1): 5–7.

Derakhshan, A. and H. Khodabakhshzadeh. 2011. Why CALL why not MALL: An in-depth review of text-message vocabulary. *Learning Theory and Practice in Language Studies* 1(9): 1150–1159.

Doyle, A. 2001. Web animation: Learning in motion. *Tech & Learning* 22(2): 30–42. http://search.proquest.com/docview/212089737?accountid=48462 (accessed 26 January 2012).

Dunkel, P. 1990. Implications of the CAI effectiveness research for limited English proficient learners. *Computers in the Schools* 7(1–2): 31–52.

Education on the Go with iPhone 3GS from Maxis. 2009. http://www.maxis.com.my/mmc/index.asp?fuseaction=press.view&recID=397 (accessed 26 January 2012).

Engwall, O. and O. Bälter. 2007. Pronunciation feedback from real and virtual language teachers. *Journal of Computer Assisted Language Learning* 20(3): 235–262.

Fraser, H. 2000. Coordinating improvements in pronunciation teaching for adult learners of English as a second language. Canberra: Department of Education, Training and Youth Affairs.

Fotouhi-Ghazvini, F., R. A. Earnshaw and L. Haji-Esmaeili. 2009. Mobile assisted language learning in a developing country context. Proceedings of the 2009 International Conference on CyberWorlds, 391–397.

Gamper, J. and J. Knapp. 2002. A review of intelligent CALL systems. *Computer Assisted Language Learning* 15(4): 329–342.

Gilakjani, A. P. and A. Mohammad Reza. 2011. Why is pronunciation so difficult? *Canadian Center of Science and Education* 4(3): 74–83.

Godwin-Jones, R. 2005. Messaging, gaming, peer-to-peer sharing: Language learning strategies and tools for the millennial generation. *Language Learning & Technology* 9(1): 17–22.

Graesser, A. C., P. Chipman and B. G. King. 2008. Computer-mediated technologies. In *Handbook of research on educational communications and technology*, 3rd ed., eds. J. M. Spector, M. D. Merrill, J. J. G. van Merriënboer and M. Driscoll, 211–224. New York: Lawrence Erlbaum Associates.

Herrington, J., S. McKenzie, R. Pascoe, A. Woods-McConney, J. MacCallum and P. Wright. 2010. A whole-school approach to technological literacy: Mobile learning and the iPhone. In *Proceedings of Global Learn Asia Pacific 2010*, eds. Zoraini Wati Abas et al., 181–186. AACE Publications.

Hult, S., M. Kalaja, O. Lassila and T. Lehtisalo. 1990. HyperReader-An interactive course in reading comprehension. *System* 18(2): 189–198.

Jesse, A. and D. W. Massaro. 2010. The temporal distribution of information in audiovisual spoken-word identification. *Attention, Perception & Psychophysics* 72(1): 209–225, doi:10.3758/APP.72.1.209.

Johnson, W. L., J. W. Rickel and J. C. Lester. 2000. Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education* 11: 47–78.

Joseph, S. and M. Uther. 2009. Mobile devices for language learning: Multimedia approaches. *Research and practice in Technology Enhanced Learning* 4(1): 7–32.

Kadyte, V. 2004. Learning can happen anywhere: A mobile system for language learning. In *Learning with mobile devices*, eds. J. Attewell and C. Savill-Smith, 73–78. London: Learning and Skills Development Agency. http://www.lsda.org.uk/pubs/ (accessed 20 January 2012).

Kayaoğlu, M. N., R. Dağ Akbaş and Z. Öztürk. 2011. A small scale experimental study: Using animations to learn vocabulary. *Turkish Online Journal of Educational Technology* 10(2): 24–30.

Koda, T. and P. Maes. 1996. Agents with faces: The effects of personification of agents. In *Proceedings of HCI,* 98–103. The British HCI Group.

Kramsch, C. and R. Andersen. 1999. Teaching text and context through multimedia. *Language Learning and Technology* 2(2): 31–42. http://llt.msu.edu/vol2num2/article1/index.html. (accessed 24 March 2012).

Kukulska-Hulme, A. 2010. Mobile learning as a catalyst for change. *Open Learning: The Journal of Open and Distance Learning* 25(3): 181–185.

———. 2005. Mobile usability and user experience. In *Mobile learning: A handbook for educators and trainers*, eds. A. Kukulska-Hulme and J. Traxler, 45–56. London: Routledge.

Kukulska-Hulme, A. and L. Shield. 2008. An overview of mobile assisted language learning: From content delivery to supported collaboration and interaction. *ReCALL* 20(3): 271–289.

Kukulska-Hulme, A., J. Traxler and J. Pettit. 2007. Designed and user-generated activity in the mobile age. *Journal of Learning Design* 2(1): 52–65.

Lazaraton, A. 2004. Gesture and speech in the vocabulary explanations of one ESL teacher: A micro-analytic inquiry. *Language Learning* 54: 79–117.

Lester, J. and B. Stone. 1997. Increasing believability in animated pedagogical agents. In *Proceedings of the First International Conference on Autonomous Agents*, eds. Johnson, W. L. and B. Hayes–Roth, 16–21. New York, USA: ACM New York.

Lin, H., T. Chen and F. M. Dwyer. 2006. Effects of static visuals and computer-generated animations in facilitating immediate and delayed achievement in the EFL classroom. *Foreign Language Annals* 39(2): 203–219.

Liu, M., Z. Moore, L. Graham and S. Lee. 2002. A look at the research on computer-based technology use in second language learning: A review of the literature from 1990–2000. *Journal of Research on Technology in Education* 34(3): 250–273.

Liu, M. 1994. Hypermedia assisted instruction and second language learning: A semantic-network-based approach. *Computers in School* 10(3–4): 293–312.

Liu, Y., D. W. Massaro, T. H. Chen, D. Chan and C. A. Perfetti. 2007. *Using visual speech for training Chinese pronunciation: An in-vivo experiment.* http://www.cs.cmu.edu/~max/mainpage_files/files/SLaTE07_Liu_Training_Chinese_Pronunciation.pdf.

Lun, E. V. n.d. *Talking head.* http://www.chatbots.org/talking_head/ (accessed 18 December 2011).

Malik, S. and M. S. Shabbir. 2008. Perceptions of university students on self-directed learning through learning technology. *European Journal of Scientific Research,* 24(4): 1523–1537.

Massaro, D. W. 2006a. Embodied agents in language learning for children with language challenges. In *Proceedings of the 10th International Conference on Computers Helping People with Special Needs, ICCHP 2006,* eds. K. Miesenberger, J. Klaus, W. Zagler and A. Karshmer, 809–816. Berlin: Springer.

———. 2006b. The psychology and technology of talking heads: Applications in language learning. In *Natural, intelligent and effective interaction in multimodal dialogue systems,* eds. O. Bernsen, L. Dybkjaer and J. van Kuppevelt, 183–214. Dordrecht, The Netherlands: Kluwer Academic Publishers.

———. 1998. *Perceiving talking faces: From speech perception to a behavioural principle.* Cambridge, MA: MIT Press.

Mayer, R. and R. Moreno. 2002. Animation as an aid to multimedia learning. *Educational Psychology Review* 14(1): 87–99.

———. 1998. A cognitive theory of multimedia learning: Implications for design principles. Paper presented at the annual meeting of the ACM SIGCHI Conference on Human Factors in Computing Systems. Los Angeles, CA.

McMenemy, K. and S. Ferguson. 2009. Enhancing the teaching of professional practice and key skills in engineering through the use of computer animation. *International Journal of Electrical Engineering Education* 46(2): 164–174. http://search.proquest.com/docview/219870653?accountid=48462 (accessed 26 January 2012).

Mohd Muzhafar Idrus. 2011. Impact of wireless technology on language learning. Paper presented at the WVTESOL 14th Annual Spring Conference, Ripley, West Virginia.

Moreno, R. and R. Mayer. 1999. Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology* 91(2): 358–368.

Moreno, R. and R. E. Mayer. 2000. Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology and Technology* 92: 724–733.

Motteram, G. and G. Stanley. 2011. Web 2.0 and the normalisation of CALL. *International Journal of Computer Assisted Language Learning and Teaching* 1(2): 1–6.

Park, C. N. and J. Son. 2009. Implementing computer-assisted language learning in the EFL classroom: Teachers perceptions and perspectives. *International Journal of Pedagogies & Learning* 5(2): 80–101.

Pettit, J. and A. Kukulska-Hulme. 2007. Going with the grain: Mobile devices in practice. *Australasian Journal of Educational Technology* (AJET) 23(1): 17–33.

Picard, R. W. 1997. *Affective computing*. Cambridge, MA: MIT Press.

Pronunciation. 2011. Melbourne, AU: University of Melbourne

Rodgers, T. 2001. Language teaching methodology. www.cal.org/resources/Digest/rodgers.html (accessed 30 September 2011).

Saba, A. 2009. *Benefits of technology integration in education*. http://edtech2.boisestate.edu/sabaa/502/Saba_Synthesis_Paper.pdf (accessed 5 June 2013).

Saran, M., G. Seferoglu and K. Cagiltay. 2009. Mobile assisted language learning: English pronunciation at learners' fingertips. *Eurasian Journal of Educational Research* 34: 97–114.

Schnotz, W. and T. Rasch. 2005. Enabling, facilitating and inhibiting effects of animations in multimedia learning: Why reduction of cognitive load can have negative results on learning. *Educational Technology Research and Development* 53(3): 47–58.

Sharples, M., J. Taylor and G. Vavoula. 2007. A theory of learning for the mobile age. In *The Sage handbook of e-learning research,* eds. R. Andrews and C. Haythornthwaite 221–247. London: Sage Publications.

———. 2005. Towards a theory of mobile learning. Paper presented at 4th World Conference on mLearning, Cape Town, South Africa. http://www.mlearn.org.za/CD/papers/Sharples-%20Theory%20of%20 Mobile.pdf (accessed 25 May 2012).

Shaw, E., W. L. Johnson and R. Ganeshan. 1999. Pedagogical agents on the web. *Proceedings of the Third International Conference on Autonomous Agents, Seattle, WA*. http://www.isi.edu/isd/ADE/papers/agents99/agents99.htm (accessed 15 March 2012).

Sime, D. 2006. What do learners make of teachers' gestures in the language classroom? *International Review of Applied Linguistics in Language Teaching* 44(2): 211–230.

Sood, K. 2010. Mobile media archive using significant tools for mobile assisted language learning (PhD diss., San Diego State University) http://hdl.handle.net/10211.10/602 (accessed 11 November 2011).

Stiller, K. D., A. Freitag, P. Zinnbauer and C. Freitag. 2009. How pacing of multimedia instructions can influence modality effects: A case of superiority of visual texts. *Australasian Journal of Educational Technology* 25(2): 184–203.

Stockwell, G. 2010. Using mobile phones for vocabulary activities. Examining the effect of the platform. *Language Learning & Technology* 14(2): 95–110.

Sumby, W. H. and I. Pollack. 1954. Visual contribution to speech intelligibility in noise. *Journal of the Acoustical Society of America* 26 (2): 212–215.

Sundberg P. A. 1998. Animation in CALL: Learning to think in the fourth dimension. Paper presented at CALICO '98 Symposium San Diego, California.

Tamburini, F. and S. Paci. 2002. Web based language learning: Authoring and assessment technologies. International Conference on Information and Communication Technology in Education, Badajoz, Spain.

Tversky, B., J. B. Morrison and M. Bétrancourt. 2002. Animation: Can it facilitate? *International Journal of Human-Computer Studies* 57(4): 247–262.

Valk, J., A. Rashid and L. Elder. 2010. Using mobile phones to improve educational outcomes: An analysis of evidence from Asia. *International Review of Research in Open and Distance Learning* 11(1): 1492–3831.

van Mulken, S. A. and E. Andre. 1998. The Persona effect: How substantial is it? In *Proceedings of HCI98,* 53–66.

Voce, J. and M. J. Hamel. 2001. Towards a compromise between talking heads & interface agents: A web-based "mentor" for computer assisted language learning (CALL). In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, eds. C. Montgomerie and J. Viteli, 1964–1965. Chesapeake, VA: AACE.

Walker, J. H., L. Sproull and R. Subramani. 1994. Using a human face in an interface. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Celebrating interdependence,* 85–91.

Williamson, V. M. and M. R. Abraham. 1995. The effects of computer animation on the particulate mental models of college chemistry students. *Journal of Research in Science Teaching* 32(5): 521–534.

Wik, P. 2011. The virtual language teacher: Models and applications for language learning using embodied conversational agents. PhD diss., KTH Royal Institute of Technology, Sweden.

Wik, P. and A. Hjalmarsson. 2009. Embodied conversational agents in computer assisted language learning. *Speech communication* 51(10): 1024–1037.

Xiao, X. and M. G. Jones. 1995. Computer animation for EFL learning environments. Paper presented at the Annual Conference of the International Visual Literacy Association, Chicago, IL.