

RESEARCH ARTICLE

Open Access



# Does a pedagogical agent's gesture frequency assist advanced foreign language users with learning declarative knowledge?

Robert O. Davis<sup>1\*</sup> , Joseph Vincent<sup>2</sup> and Lili Wan<sup>3</sup>

\*Correspondence:

red1020@gmail.com

<sup>1</sup> Department of English  
Linguistics and Language  
Technology, Hankuk  
University of Foreign Studies,  
107 Imun-ro, Imun-dong,  
Dongdaemun-gu, Seoul,  
South Korea

Full list of author information  
is available at the end of the  
article

## Abstract

Since the conception of pedagogical agents in multimedia environments, researchers have advocated for agents to be designed to exhibit social cues that prime the social interaction of the target audience. One powerful social cue has been agent gesturing. While most agents are created only to use deictic (pointing) gestures, there is recent evidence that agents that perform all gesture types (iconic, metaphoric, deictic, and beat) with enhanced frequency help foreign language users learn more procedural knowledge. Therefore, this research examines how all gesture types and different frequencies influence agent persona and learning outcomes when foreign language users learn declarative knowledge. The results indicated that the use of gestures, regardless of frequency, significantly increase agent persona. However, gesture frequency produced conflicting learning outcomes. While enhancing gestures were beneficial for cued recall and recognition, the average gesture condition was not, which indicates that the strength of social cues is important.

**Keywords:** Pedagogical agents, Agent persona, Embodiment principle, Cued recall

## Introduction

After two decades of research on pedagogical agents in multimedia environments, meta-analyses and reviews provide evidence that agent presence is beneficial for learning outcomes (Davis 2018; Schroeder et al. 2013; Wang et al. 2017). However, the physical presence of the agent must also provide social cues found in human-to-human communications. *Social agency theory*, one of the early theories examining social cues, suggested that the image and voice activate social interaction schemas which allow for deeper cognitive processing by the users (Mayer et al. 2003). While early research focused more on how voice (human vs. computer) impacted agent perception and learning, later research began to explore the social cuing of the agent's image. Mayer and Dapra (2012) examined the nonverbal communicative aspects of an agent with the *embodiment principle*, which proposes that agents that exhibit more human-like characteristics, such as eye-gaze, facial expressions, body sway, and gestures, help users learn more deeply when compared to conditions without embodiment. During the course of three experiments, the authors found that agents exhibiting high embodiment, the use

of the aforementioned nonverbal cues, significantly improved agent perception and the transfer of learning when measured against agents with no (low) embodiment features. However, the concept of high embodiment still requires further development as certain embodiment features have minimally addressed the potential of nonverbal cueing capabilities. For example, a meta-analysis on agent gesturing supported the embodiment principle by finding significantly higher agent persona and learning outcomes (retention and near transfer) than conditions without embodiment (Davis 2018), but one of the limitations of the analysis was that all agents used deictic (pointing) gestures. This limitation might be significant in terms of high embodiment and activating social schema because it is only one of the four potential gesture types people use to nonverbally convey information when communicating. Recent studies exploring gestures have indicated that multiple gesture types can be helpful for certain persona subscales, but learning outcomes can be more complicated (Davis and Antonenko 2017; Davis and Vincent 2019). One aspect that is gaining more attention, and might play an important role, is the type of information that is being learned. Davis and Antonenko (2017) found no significant differences with gestures when Korean fifth- and sixth-grade students learned English grammar. However, Korean university students learning procedural knowledge in the scientific domain indicated that gestures and frequency could assist with higher learning outcomes (Davis and Vincent 2019). Another study that did not use gestures, but examined an agent's assistance with learning declarative knowledge, indicated that an agent using eye gaze helped students learn more declarative knowledge (Fountoukidou et al. 2019). Therefore, the type of knowledge domain might be an important component to learning outcomes with embodied agents. The purpose of this study is to examine how gesture frequency (enhanced vs. average vs. no) affects agent persona and the learning outcomes of cued recall and recognition with declarative knowledge with advanced university foreign language users.

## Literature review

### Gestures

In the course of human-to-human nonverbal communication, people perform a mixture of iconic, metaphoric, deictic, and beat gestures. Iconic gestures represent concrete information pertaining to objects like the “sun,” or a “mirror;” while metaphoric gestures illustrate abstract information such as “intelligence” or “evil.” Deictic gestures (pointing) direct spatial attention, and beat gestures are movements that are perfectly timed to emphasize important words or phrases with the intensity of the speech (Theune and Brandhorst 2009). These gesture types can also be classified according to the information they present. Iconic, metaphoric, and deictic gestures are considered representational gestures because they communicate or support the information contained in speech, whereas beat gestures are classified as nonrepresentational because they do not communicate semantic information (Hostetter 2011). See Fig. 1 for examples of representational gestures.

Kelly et al. (2009) provides an example of how representational gestures can illustrate a driver involved in a car accident being surprised by an unseen car hitting them at an intersection. While explaining the beginning sequences of the collision, the person performs a gesture with one hand representing their car traveling through the intersection



and the other hand representing the other car crashing into their own car's back passenger side, which would explain why the driver was surprised. The two cars, driving through the intersection, and the collision of the two cars would be considered representational gestures. If the speaker wanted to convey more information, he or she could perform hand gestures at a faster rate to denote the speed of the vehicles, or hit the hands together more forcefully to illustrate the intensity of the impact. Therefore, representational gestures can be redundant with spoken information, or non-redundant by conveying unspoken information. Likewise, nonrepresentational gestures, such as beat gestures, could be used by moving the hands up and down with a phrase such as "out of nowhere" to emphasize the surprise. Thus, gestures can support information which is present within the speech being used, but also to supply additional information that could provide additional details not found in the spoken words.

For this reason, gestures are considered another form of language production (Kendon 2004); and gestures and speech could be seen as an integrated system since gestures without speech are incomprehensible (McNeill 1992), and speech absent of gestures is more difficult to understand (Krauss et al. 1991). Various brain-imaging techniques have found support for the integrated system and the ability to discern information from gestures. Research using fMRI technology showed that the Broca's area, better known as the language center of the brain, processes gestures performed in speech (Willems and Hagoort 2007), and PET scans indicate that the Broca's area activates when viewing hand and arm movements (Schlaug et al. 1994). However, the brain is an efficient processor and does not haphazardly activate. Further PET research has indicated that the Broca's area does not register random movements (Johnson-Frey et al. 2003), nor does it process meaningless gestures (Grèzes et al. 1999).

### **Gestures and foreign language**

The extra processing ability of the brain to comprehend speech and gestures is beneficial for foreign language users listening to non-native speech. Since non-native speakers often lack the verbal proficiency of native speakers, it has been suggested that gestures are an important feature to assist foreign language users with the ability to communicate and comprehend speech (McCafferty 2002). This might be due to the listening strategies used by each group. Native speakers normally employ a top-down approach that focuses on listening for the main ideas, whereas non-native speakers implement a bottom-up approach (Griffiths 1990), especially during the earlier levels of language learning. Church et al. (2004) found that non-English speaking Spanish students learned mathematical concepts presented in English with gestures at significantly higher rates than students in a speech-only condition. Even though the Spanish speaking students were not able to comprehend the language, the authors concluded that the gestures allowed the learners to access their lexical understanding of Spanish to help them learn the concepts presented in English. Likewise, Kelly et al. (2009) discovered that gestures helped English-speaking American university students learn significantly more Japanese verbs, and the results were maintained in follow up tests one and 2 weeks later.

To better understand how gestures help with comprehension, Hostetter (2011) conducted a meta-analysis across different populations and provided effect sizes using Cohen's *d*. Regarding the use of gestures for comprehension in populations with less proficient verbal skills such as non-native speakers, Hostetter found that gestures benefitted non-native speakers ( $d=0.85$ ), but when compared against native speaker samples ( $d=0.60$ ), these benefits were not significant. Even though the results between non-native and native populations were not significant, the large effect sizes found in the non-native analysis support McNeill's (1992) suggestion that gestures and speech work as one modality to assist comprehension.

### **Gestures and pedagogical agents**

Although gestures have been found to be helpful for learning in human-to-human interaction with non-native speakers, the results have been mixed with gesturing pedagogical agents in multimedia environments. Experiments examining non-native English-speaking university students learning English grammar with agents using deictic gestures have shown no significant differences in learning (Carlotto and Jaques 2016; Choi and Clark 2006). Likewise, Korean elementary fifth and sixth graders failed to show significance in retention and transfer tests on English grammar when viewing agents which used all gesture types (full gesture) against an agent that only used deictic gestures and an agent that did not gesture. However, the full gesture condition was perceived as more human-like and engaging than the no-agent condition in agent persona. Conversely, all these studies had a common denominator—they focused on English Grammar. Larsen-Freeman (2003) suggests that teaching grammar could be problematic because rule-based decontextualized grammar is not quickly learned, but requires experience and practice which allows the learner to exceed the boundaries of context. In other words, the short-term exposure of little known grammar concepts might limit the potential to find significance between conditions.

On the other hand, when researchers have focused on non-grammatical aspects in a foreign language, the results have suggested that gestures impact learning outcomes. Yung and Paas (2015) found that 7th grade Taiwanese students learned significantly more about the cardiovascular system with a virtual human that performed deictic gestures than in an absent-agent condition. Likewise, Davis and Vincent (2019) examined the use of gesture type and frequency with foreign language university students learning procedural information. Using Alibali et al. (2001) average number of representational and nonrepresentational gestures when humans talk about procedural information, the research included: an average gesture condition; an enhanced gesture condition (which doubled the average gesture rate); a conversational gesture condition that only performed movements which carried no semantic meaning; and a no-gesture condition. The results indicated that the enhanced gesture condition scored significantly higher than the no-gesture condition and approached significance against the conversational gesture condition. Since there were no other significant conditions, this research showed that increasing the gesture rate could be beneficial to foreign language users' ability to freely recall procedural information. Therefore, these studies suggest that gestures can play an important role in learning, but that the type of information being taught could affect learning outcomes.

### **Knowledge domains**

In the research literature examining pedagogical agents and their impacts on the type of information being learned, meta-analyses have been more concerned with instructional domain than knowledge domain. Using the effect size formula of Hedges'  $g$ , Schroeder and Adesope (2013) found that agents in science ( $g=0.27$ ) and math ( $g=0.28$ ) domains had significantly better learning outcomes, while those in the humanities ( $g=0.06$ ) did not significantly increase learning when compared against control conditions with no agent present. Likewise, Davis (2018) found that gesturing agents had higher effect sizes when measuring transfer within instructional domains of science ( $g=0.47$ ) and math ( $g=0.32$ ) when compared to effect size in the humanities ( $g=0.08$ ) with static or no agent conditions. While these instructional domains do not suggest a type of knowledge being learned, research within the field has only recently started to account for the type of knowledge being taught and learned with pedagogical agents.

Regarding pedagogical agent research that has focused on knowledge domain, the aforementioned research by Davis and Vincent (2019) with various gesture types and gesture frequencies for foreign language users learning procedural knowledge, and an experiment examining computer training and the recall of declarative knowledge (Fountoukidou et al. 2019). The Fountoukidou et al. (2019) research tested an embodied agent with eye gaze and body movements, and no facial expressions or gestures with arms or hands, with learning to perform Internet searches using only the agent's eyes and then testing the recall of declarative knowledge. The results of the study indicate that the agent with limited embodiment assisted users with recalling more declarative knowledge.

Of the two knowledge domains tested with pedagogical agents, declarative knowledge refers to the facts, concepts, and generalizations of what is; and procedural knowledge, which requires specific steps to be understood (Marzano 1997). Researchers have

determined that these two types of information are separately stored and activated in the brain. The hippocampus and temporal cortex store declarative knowledge, while procedural knowledge resides in the basal ganglia and the frontal cortex (Ullman 2016). Even though these knowledge types are distinct and are processed in different areas of the brain, they do interact on some levels since the learner has to process “what is” (declarative knowledge) in order to comprehend specific steps (procedural knowledge) of the information. However, different teaching strategies have been suggested for learning each knowledge type. Marzano (1997) suggest three distinct phases for learning declarative knowledge: constructing meaning, organizing, and the storing of information. One of the keys for storing declarative information is for learners to see physical or pictographic representations so they can create mental pictures of information (Brown, 1995, Marzano 1997). On the other hand, the three phases of learning procedural information are: constructing models, shaping, and internalizing the information (Marzano 1997). For students to learn procedural information, the information must be presented in well-defined steps so learners can work out the process.

Therefore, this research attempts to examine whether social cues (gestures) affect foreign language users in their comprehension of declarative knowledge. Previous researchers have found that gesture frequency is effective in helping foreign language learners comprehend procedural knowledge; but gesture frequency and declarative knowledge outside of grammar rules and concepts have not been tested with foreign language users. The testing of gesture frequency is important because education requires multiple types of knowledge to be learned.

### Research questions

RQ1: To what extent does gesture frequency (enhanced vs. average vs. no) affect foreign language users’ perception of agent persona when learning declarative knowledge?

RQ2: To what extent does gesture frequency (enhanced vs. average vs. no-gesture) affect foreign language users’ learning outcomes of cued recall and recognition when learning declarative knowledge?

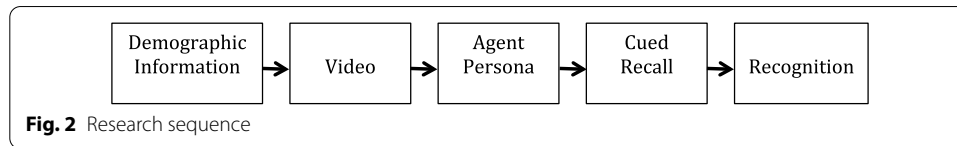
## Methods

### Research design and participants

This research used a between-subjects experimental design to test whether the gesture frequency (enhanced vs. average vs. no-gesture) affected foreign language users’ social perception (agent persona) of the agent, and whether gesture frequency assisted with the recognition and cued recall of declarative information. Participants were randomly assigned to each condition. The research sequence can be seen in Fig. 2. In all, the duration of the experiment was not more than 35 min.

The participants for this experiment were 154 foreign language students at a university in Seoul, South Korea. According to demographic data collected before the experiment, 56% (87) of the participants were Korean, while 44% (67) of the participants were from countries such as China, Germany, Uzbekistan, Ghana, Russia, Ukraine, Egypt, Pakistan, Nepal, Japan, Indonesia, and Vietnam. The average age of the participants was 22.4 years ( $SD=2.8$ ) and all participants were first or second majors in an English-focused major,





and taking classes delivered in English. Of the 154 participants, 59 identified as male, 92 as female, and three chose not to answer.

### Multimedia environment

The video environment was created in iClone™ 7.1 with the character Mason in an open grassy field with rock formations in the background. To test how gesture frequency affects the agent persona and learning outcomes with declarative knowledge, the video did not include verbal redundancy or visual aids. Only the background and agent were present on the screen. See Fig. 3 for an image of the agent and the background. The agent in the experimental conditions (enhanced, average) included facial expressions, lip synchronization, eye gaze, small instances of body sway, and gestures. The agent in the control condition was a static agent with lip synchronization. The total video viewing time was 6 mins and 6 s.

### Gesture design

The gestures performed by the agents were created in the iClone™ 7.1 timeline with the character puppet feature. The no-gesture condition agent kept hands together in front of the stomach, while the gesture conditions consisted of iconic, metaphoric, deictic, and beat gestures. To set the average rate for gestures, this experiment used the average gesture rates of 6.51 representational gestures and 1.32 nonrepresentational gestures (beat) per 100 words (Hostetter and Skirving 2011). The average gesture condition performed 6.49 representational gestures (iconic: 21, metaphoric: 30, deictic: 4) and 1.29 nonrepresentational gestures (beat: 11), while the enhanced gesture condition performed double that amount with 13.1 representational gestures (iconic: 36, metaphoric: 70, deictic: 5) and 2.6 nonrepresentational gestures (beat: 22) per 100 words.



**Fig. 3** Gesturing agent explaining information about Australia

The script was divided into eight 100-word sections and an additional section of 47 words. The number of gestures scripted and performed was based off 100-word sections, so in the average condition every 100 words would have five to seven gestures, while the enhanced condition would contain 13 or 14 gestures per 100 words. In addition, when information focused on answers that would be in the cued recall and recognition tests, gestures for that information were the same in both conditions. The only difference was that the enhanced gesture condition included more gestures between key information points than the average gesture condition.

### **Instructional content**

The content for this experiment, *Discovering Australia*, has previously been used in several studies (Ritzhaupt and Barron 2008; Ritzhaupt and Kealy 2015; Ritzhaupt et al. 2011; Wang et al. 2018). The topic consists of 22 separate bodies of text that contain 74–80 words in each; but, two bodies of text serve as introductions that have no questions associated with the information. The other 20 bodies of text contained two questions each. For this research, one introduction body and ten bodies of text were chosen. The topics were: Morning Glory clouds, northern coastal wetlands, Aboriginals, Perth, East Timor Sea, animals in Australia waters, wild rabbits, MacDonnell Ranges, Wolf Creek Crater, and the Great Sandy Desert. The Flesch-Kincaid grade level for the whole body of text was 11.2, which would be considered 11th grade material.

### **Instruments**

#### **Demographic data**

Participants were required to answer demographic data questions to establish their international age (Koreans use a different system), gender, year in university, native language, major, and whether they had lived in or visited Australia for an extended period of time. Those who had lived or spent more than six months in Australia were removed from the pool. Also, participants were asked to contact the monitor and abort the experiment if they had previously participated in this research. Since this research topic has primarily been conducted in America, and this was the first time using this content in South Korea, there were no instances of participants having previous exposure this research topic. The data was used to assess whether any significant demographic items influenced the results.

#### **Prior knowledge test**

Previous experiments which utilized this content did not have a prior knowledge test, since American students were deemed to not have significant knowledge to affect the results due to geographical distance. However, since Australia study abroad programs are well known in Korea, establishing the degree of prior knowledge was a necessity for this research. Before watching the video on Australia, participants were asked to rate their prior knowledge of Australia's history, geography, and wildlife on a five-point Likert scale with (1) being "very little" and (5) being "very much." General questions such as, "How much knowledge do you have about Australia's wildlife," were asked to avoid any testing effect scenarios. Prior knowledge scores were assessed using a one-way analysis of variance (ANOVA) test to evaluate whether there were any significant differences between conditions. The results



**Table 1** Means and standard deviations for agent persona subscales, recognition, and cued recall

Measure	Enhanced gestures (n = 52)		Average gestures (n = 51)		No gestures (n = 51)	
	M	SD	M	SD	M	SD
Facilitation	29.06	8.35	27.55	8.00	20.35	7.81
Credibility	17.00	3.45	16.45	3.92	13.84	3.81
Human-likeness	12.19	4.02	12.12	3.37	8.84	3.23
Engagement	13.46	4.09	13.10	4.17	9.33	3.26
Recognition	10.06	3.87	8.59	3.84	8.45	3.22
Cued Recall	3.73	3.30	2.55	3.15	2.41	2.55

indicated there were no significant differences between conditions ( $F(2, 151) = 1.272$ ,  $p = 0.283$ ).

### Agent persona

Since gestures have been shown to increase the perception of agent persona, this experiment seeks to test whether gesture frequency impacts the learner's perception of the agent. The agent persona instrument revised (API-R; Schroeder et al. 2018) was used to evaluate agent persona. The API-R consists of 25 Likert scale questions ranging from 1 (strongly disagree) to 5 (strongly agree) within the four subscales of facilitation ( $\alpha = 0.94$ ), credibility ( $\alpha = 0.81$ ), human-likeness ( $\alpha = 0.80$ ), and engagement ( $\alpha = 0.87$ ). The facilitation subscale consists of 10 questions, while credibility, human-likeness, and engagement each contain five questions.

### Cued recall test

Before participants could answer multiple-choice recognition questions, they responded to two cued recall questions from each of the ten topics covered in the presentation. Following are two sample questions: (1) How is the rabbit disease myxomatosis transmitted? and (2) What relatively harmless reptile inhabits the coastal wetlands? Answers were evaluated by two independent raters. Each correct answer was worth one point. Cohen's Kappa was used to measure inter-rater reliability using SPSS 25. The results showed that the two raters had substantial agreement with a reliability score of  $k = 0.69$  (Cohen 1960). Disagreements in scoring were reconciled by discussions between the two raters.

### Recognition test

Immediately after finishing the cued recall section, participants took a multiple choice recognition test. Each correct answer was worth one point. An example question is: "What techniques have anthropologists used to determine the time the aborigines came to Australia?" The internal consistency was measured at  $\alpha = 0.70$ , which is considered acceptable. See Table 1 for descriptive statistics for all measures.

### Estimation approach

In the experiment, the dependent variables of persona, cued recall, and recognition were measured with multiple correlated items for each participant. Moreover, independent

variables and control variables used in the model are all time invariant variables, which cannot be estimated in fixed effect models. Each participant was randomly assigned to one of the treatment groups or the control group, thus unobserved individual heterogeneities do not exist in model estimation. As a result, a random-effect linear regression model was employed to evaluate the persona of the agent, and a random-effect logit regression model was used to evaluate cued recall and recognition.

Persona was measured according to the subscales of facilitation, credibility, human-likeness, and engagement. The random-effect linear regression model is used to evaluate each subscale respectively. Specifically, the estimated model for agent persona was:

$$\text{Persona subscale}_{ij} = \alpha + \beta \times \text{Gesture}_i + \zeta \times X_i + \theta_i + \varepsilon_{ij}$$

$\text{Persona subscale}_{ij}$  measured the participant's evaluation of the agent for each of the persona subscales. Each dimension was measured with multiple questions, where  $i$  represents participant,  $j$  denotes question.  $\text{Gesture}_i$  was a set of dummy variables which indicated whether participant  $i$  watched a video with enhanced gestures, average gestures, or no-gestures.  $X_i$  was a set of control variables for participant-specific characteristics, such as *age*, *school year*, *gender*, *English major* (whether English is the primary or secondary major), and *prior knowledge* about Australia's history, geography, and wildlife.  $\theta_i$  was the participant-specific random effect, and  $\varepsilon_{ij}$  is the error term.

Cued recall and recognition were measured via multiple questions with binary codes. A random-effect logistic regression model was used to estimate the likelihood of recall or recognition separately for each participant.

$$\text{Recall(orRecognition)likelihood}_{ij} = \frac{e^{\alpha + \beta \times \text{Gesture}_i + \zeta \times X_i + \theta_i + \varepsilon_{ij}}}{1 + e^{\alpha + \beta \times \text{Gesture}_i + \zeta \times X_i + \theta_i + \varepsilon_{ij}}}$$

where  $i$  represents each participant,  $j$  denotes each recall or recognition question,  $\text{Recall(orRecognition)likelihood}_{ij}$  represents the probability of the recall or recognition of participant  $i$  for question  $j$ . Other variables used in the model are the same as for the aforementioned linear regression model used to estimate persona.

## Results

Table 2 presents results for each subscale of agent persona (facilitation, credibility, human-likeness, engagement). There were three gesture conditions (enhanced, average, and no-gesture) in which to compare differences. The enhanced and average gesture conditions were used as the base conditions respectively. Models (1), (3), (5), and (7) are regression results which used enhanced gestures as a base condition, and models (2), (4), (6), and (8) are average gestures as a base condition. From models (1) to (8), the evaluation of the no-gesture condition was significantly lower than the enhanced gesture and average gesture conditions for all of the agent persona subscales. However, there are no significant differences between the average gesture and enhanced gesture conditions. Specifically, the evaluation of facilitation for the no-gesture condition was 0.832 ( $p < 0.01$ ) points lower than the enhanced gesture condition, and 0.693 ( $p < 0.01$ ) points lower than the average gesture condition. For credibility, the evaluation of the no-gesture condition was 0.626 ( $p < 0.01$ ) points lower than the enhanced gesture condition and

**Table 2** Regression results for agent persona

	Facilitation		Credibility		Human-likeness		Engagement	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Enhanced gesture		0.140 (0.158)		0.087 (0.147)		0.033 (0.137)		0.056 (0.151)
Average gesture	-0.140 (0.158)		-0.087 (0.147)		-0.033 (0.137)		-0.056 (0.151)	
No gesture	-0.832** (0.159)	-0.693** (0.159)	-0.626** (0.148)	-0.539** (0.149)	-0.731** (0.138)	-0.698** (0.138)	-0.835** (0.152)	-0.779** (0.153)
Age	-0.029 (0.029)	-0.029 (0.029)	-0.045 (0.027)	-0.045 (0.027)	0.017 (0.025)	0.017 (0.025)	-0.013 (0.028)	-0.013 (0.028)
Year	-0.097 (0.068)	-0.097 (0.068)	-0.041 (0.063)	-0.041 (0.063)	-0.135* (0.059)	-0.135* (0.059)	-0.118 (0.065)	-0.118 (0.065)
Gender	-0.122 (0.149)	-0.122 (0.149)	-0.005 (0.138)	-0.005 (0.138)	-0.123 (0.129)	-0.123 (0.129)	-0.269 (0.142)	-0.269 (0.142)
English major	-0.079 (0.148)	-0.079 (0.148)	0.065 (0.138)	0.065 (0.138)	0.037 (0.128)	0.037 (0.128)	0.214 (0.141)	0.214 (0.141)
Prior knowledge 1	0.202 (0.136)	0.202 (0.136)	0.220 (0.127)	0.220 (0.127)	0.312** (0.118)	0.312** (0.118)	0.186 (0.130)	0.186 (0.130)
Prior knowledge 2	0.020 (0.114)	0.020 (0.114)	-0.090 (0.106)	-0.090 (0.106)	-0.146 (0.099)	-0.146 (0.099)	-0.052 (0.109)	-0.052 (0.109)
Prior knowledge 3	0.056 (0.106)	0.056 (0.106)	-0.022 (0.099)	-0.022 (0.099)	-0.170 (0.092)	-0.170 (0.092)	-0.025 (0.102)	-0.025 (0.102)
Constant	3.477** (0.659)	3.337** (0.667)	4.327** (0.614)	4.240** (0.622)	2.605** (0.572)	2.572** (0.579)	3.169** (0.631)	3.113** (0.639)
N	1540	1540	770	770	770	770	770	770

Standard errors in parentheses, \*  $p < 0.05$ , \*\*  $p < 0.01$

0.539 ( $p < 0.01$ ) points lower than the average gesture condition. Similarly, in human-likeness, the no-gesture condition was 0.731 ( $p < 0.01$ ) points lower than the enhanced gesture condition and 0.698 ( $p < 0.01$ ) points lower than the average gesture condition. Finally, in engagement, the no-gesture condition was 0.835 ( $p < 0.01$ ) points less than the enhanced gesture condition and 0.779 ( $p < 0.01$ ) points less than the average gesture condition. Using gestures together to explain the content in the video resulted in higher evaluations for agent persona. While the enhanced gesture condition performed better than the average gesture condition, the difference was not significant.

The regression coefficient indicated that the control variable *school year* had a negative effect only for human-likeness ( $b = -0.135$ ,  $p < 0.05$ ). Students with higher school years (first- through fourth-year in university) perceived the agent as less human-like. Furthermore, research question 1 of the prior knowledge assessment had a positive effect on human-likeness ( $b = 0.312$ ,  $p < 0.01$ ). More prior knowledge of Australian history led to a higher human-likeness evaluation. Other control variables were not significant.

Cued Recall and recognition were tested using a series of logistic regression models. Similarly, models (1) and (3) used the enhanced gesture condition as the base condition, while models (2) and (4) used the average gesture condition as the base condition. In model (1), participants in the average gesture condition showed a lower ability

for cued recall than those in the enhanced gesture condition ( $b = -0.310, p < 0.05$ ). Participants in the no-gesture condition had a lower possibility of cued recall than those in the enhanced gesture condition ( $b = -0.245, p < 0.10$ ), but the difference was not significant. In model (3) for recognition, participants in both the average gesture condition ( $b = -0.200, p < 0.05$ ) and the no-gesture condition ( $b = -0.167, p < 0.05$ ) had a significantly lower probability for recognition than those in the enhanced gesture condition. The results suggest that although enhanced gesture and average gesture are both effective in persona evaluation, only the enhanced gesture condition was effective in promoting the learning effect, while the average gesture failed to be effective.

School year positively affects recognition probability ( $b = 0.108, p < 0.01$ ), and prior knowledge about Australia’s wildlife increases both recall ( $b = 0.309, p < 0.01$ ) and recognition probability ( $b = 0.153, p < 0.01$ ). Other factors were not significant. As the school year increases and students had more knowledge about Australia’s wildlife, the learning effect was better. See Table 3 for the cued recall and recognition results.

**Discussion**

This study investigated whether and to what extent: (1) gesture frequency affected the agent’s perceived persona, and whether (2) gesture frequency during the presentation of propositional knowledge altered the cued recall and recognition of information. While previous research indicated that increasing gesture frequency significantly aided the recall of procedural information (Davis and Vincent 2019), this research examined the degree, if any, to which increasing gesture frequency benefited learners with regard to declarative knowledge.

RQ1: To what extent does gesture frequency (enhanced vs. average vs. no) affect foreign language users’ perception of agent persona when learning declarative knowledge?

The results showed that both gesture conditions, enhanced and average, significantly affected the participants’ perception of agent persona for all four subscales when

**Table 3** Logistic regression results for cued recall and recognition

	Cued recall		Recognition	
	(1)	(2)	(3)	(4)
Enhanced gesture		0.310* (0.134)		0.200* (0.084)
Average gesture	-0.310* (0.134)		-0.200* (0.084)	
No gesture	-0.245 (0.135)	0.065 (0.140)	-0.167* (0.084)	0.033 (0.084)
Age	-0.004 (0.025)	-0.004 (0.025)	0.006 (0.015)	0.006 (0.015)
Year	0.103 (0.058)	0.103 (0.058)	0.108** (0.036)	0.108** (0.036)
Gender	-0.162 (0.128)	-0.162 (0.128)	-0.131 (0.078)	-0.131 (0.078)
English major	0.214 (0.128)	0.214 (0.128)	0.135 (0.078)	0.135 (0.078)
Prior knowledge 1	0.084 (0.115)	0.084 (0.115)	0.049 (0.072)	0.049 (0.072)
Prior knowledge 2	0.123 (0.098)	0.123 (0.098)	0.071 (0.060)	0.071 (0.060)
Prior knowledge 3	0.309** (0.092)	0.309** (0.092)	0.153** (0.056)	0.153** (0.056)
Constant	-2.230** (0.563)	-2.541** (0.573)	-0.931** (0.348)	-1.131** (0.353)
N	3080	3080	3080	3080

Standard errors in parentheses, \*  $p < 0.05$ , \*\*  $p < 0.01$

compared against the control condition (no-gesture). These results support the claim that agents designed with high embodiment principles such as gestures can increase the social attributes of an agent, which prime the perception of a social relationship (Mayer and Dapra 2012). Although people typically anthropomorphize agents (Woo 2008), the social relationship can be disturbed if the agent fails to perform in a human-like manner (Reeves and Nass 1996). The significant persona ratings for both gesturing conditions might reflect the agents' ability to meet the social expectations normally displayed in human-to-human interaction, since gestures are a normal part of interaction between people.

However, high embodiment may not fully explain these results because experiments designed to use all gesture types and embodiment have yielded varying results when measuring agent persona (Davis and Antonenko 2017; Davis and Vincent 2019). This may be because of the participants and type of knowledge being learned. In both of the experiments, foreign language users were learning information in English, with one focusing on grammar (Davis and Antonenko 2017) and the other on procedural knowledge (Davis and Vincent 2019). Although an agent using all gesture types was rated significantly higher in human-likeness and engagement (Davis and Antonenko 2017), the content focused on English grammar, which targets rule-based extractions that require time, experience, practice, and context to master (Larsen-Freeman, 2003). In another study, Davis and Vincent (2019) focused on procedural knowledge of how lightning formed with foreign language users majoring in the humanities. The results showed that the average gesture condition was significant against the no-gesture condition in facilitation. No other conditions were significant across the other subscales or with any of the gesture conditions (enhanced, average, conversational, no-gesture). One reason for the limited findings of persona in procedural knowledge studies might be that the scientific vocabulary and information was more difficult, since procedural information requires listeners to break down information into concepts that must be logically ordered to achieve understanding of a particular goal (Willingham et al. 1989); this process with no verbal redundancy or illustrations may have consumed any attention that could be given to agent persona. However, this study on declarative knowledge about Australia may not have been as overwhelming for students in the humanities, because it allowed them more resources to analyze the agent's persona.

Although it is highly possible that gestures do increase the social acceptance of the agent, other variables might modify the significance of agent persona depending on who the participants are and the difficulty of the information being learned. Further research should examine how an agent's social cues, such as performing human-like gestures, affect the perception of an agent when measured with learner attributes and content complexity.

RQ2: To what extent does gesture frequency (enhanced vs. average vs. no-gesture) affect foreign language users' learning outcomes of cued recall and recognition when learning declarative knowledge?

The data regarding cued recall and recognition showed similar yet contradictory results. In both cued recall and recognition, the enhanced gesture condition significantly outperformed the average gesture condition in both learning outcomes, but only scored significantly higher against the no-gesture condition in recognition. There

were no other significant interactions, even though the enhanced gesture condition scored higher than the no-gesture condition ( $b = -0.245$ ,  $p < 0.10$ ) in cued recall.

Generally, these results provide some conflicting evidence for the embodiment principle (Mayer and Dapra 2012) with pedagogical agents. On one hand, the enhanced gesture condition (high embodiment) was significant against the no-gesture condition (low embodiment) with recognition, which supports the embodiment principle. On the other hand, the enhanced gesture condition failed to reach significance against the no-gesture condition with cued recall ( $b = -0.245$ ,  $p < 0.10$ ). The average gesture condition, which would be considered high embodiment, was not significant against the no-gesture condition in cued recall or recognition. However, the enhanced gesture condition (high embodiment) significantly outscored the average gesture condition (high embodiment) in both cued recall and recognition.

Although Mayer (2014) suggests that high embodiment helps participants take a social stance and learn more deeply from an agent, which is commonly suggested with tests that measure transfer, meta-analyses support the view that embodiment features produce larger effect sizes when measuring transfer scores, but lower effect sizes when measuring retention-based assessments (Davis 2018; Wang et al. 2017). However, the contradictory findings between the enhanced gesture condition and the average gesture condition suggest that test type and instant assessment could play more of a role in generating higher scores than high embodiment priming a social relationship. Delayed assessments might show that embodiment features are not as significantly different as indicated by the immediate assessment.

Likewise, the two high embodiment conditions of enhanced gestures and average gestures produced different results in learning outcomes, which may indicate the concept of high or low embodiment might be too simple. In other words, embodiment level is not some binary measure that encompasses contexts or learning materials. Since gestures are considered another form of language production (Kendon 2004), and foreign language users must mentally decipher large amounts of information while listening compared to native speakers (Carrier 1999), the concept of high or low embodiment fails to cover this specific population that could require more assistance to learn more deeply. Because of these results, researchers and designers should ask themselves what “high embodiment” means in different contexts of age, culture, and language, and assess what frequency of social cues would be beneficial for that context.

In addition, these results do support previous research that enhanced gestures can significantly increase learning outcomes (Davis and Vincent 2019); but, this study offers the first evidence that enhanced gestures significantly outperform average gestures in learning outcomes. One possible suggestion might be that the research focused on declarative knowledge, while the other study focused on procedural knowledge. Two likely reasons for the significance found in this study are that the material could be seen as less intimidating, and the agent in the enhanced gesture condition performed more gestures to scaffold the understanding of the user. Therefore, participants in the enhanced gesture condition may have been able to create more mental representations that helped them to better organize and store the information (Brown 1995; Marzano 1997). Since the enhanced and average gesture conditions used gestures for the same key information, the other gestures present in the enhanced condition may have allowed the foreign



language users more opportunity to organize the information. Whereas the average gesture condition may have provided fewer opportunities between key items of information to assist with comprehension and organizing the information.

Furthermore, following McCafferty's (2002) assessment that gestures help foreign language users to comprehend more information, the results from this study and the literature from multiple studies with pedagogical agents suggest two things: (1) the type of gestures and frequency may play a role in learning outcomes, and (2) the type of information being learned could be important for outcomes. Agents that use deictic or multiple gestures to teach grammar knowledge fail to produce any significant findings (Carlotto and Jaques 2016; Choi and Clark 2006; Davis and Antonenko 2017). Multiple gesture types and frequencies when learning procedural information can significantly increase learning outcomes, but only if the gesture frequency is enhanced (doubled) from the average frequency (Davis and Vincent 2019). When learning declarative knowledge, deictic gestures using arms (Yung and Paas 2015), multiple gestures with enhanced frequency (Davis and Vincent 2019), deictic gestures using the eyes (Fountoukidou et al. 2019), and the current study have shown significant findings against other and control conditions. It must be noted in the research using multiple gestures that those studies did not include any redundancy strategies such as keywords, which have been shown to significantly assist foreign language users with comprehension (Adesope and Nesbit 2012). Thus, this may explain why enhanced gesture frequency showed significant learning outcomes, but average gesture frequency failed to reach significance in studies for procedural knowledge (Davis and Vincent 2019) and the current study on declarative knowledge. Gesture frequency with multiple gestures and verbal redundancy will need to be examined further in future studies.

In addition, instructional designers should consider the needs of students who are studying in a second language, since higher education has seen an increase of international students studying in western countries. Although second language users are advanced enough to meet the test score criteria for admissions, this does not mean they possess the listening skills of their native speaking counterparts. Therefore, instructional designers in higher education need to consider nonverbal communication cues such as designing pedagogical agents to perform all gesture types at enhanced frequency. These strategies will be helpful for courses delivered solely online, or for courses that incorporate a flipped classroom model in which students are required to comprehend the information before attending class. Thus, course instructors and designers should consider their student population when designing online content, especially if classes include second language students.

Lastly, it is not clear why participants with increased knowledge on the topic of Australian wildlife performed better on learning measurements than the other topics. Of the 20 questions in cued recall and recognition, only four answers required the knowledge of animals. One suggestion might be that any interest in wildlife may include peripheral knowledge of habitats and other geographic information. However, there is no indication that wildlife questions statistically benefited a specific condition.

### Theoretical implications

This research does provide evidence for the embodiment principle, but raises questions about the current classifications of embodiment. While previous software limited researcher and designer ability to create human-like features such as multiple gesture types, hence the reason for only deictic gestures, current software enables designers to create agents that are more similar to humans in form and function. Thus, if the agent is to prime the participant for a social relationship that is commonly found in human-to-human interaction, then agents should be designed to human expectations. In normal interaction, people use more than one gesture, and if someone used only a pointing gesture during conversation, it would be considered socially abnormal.

Therefore, the idea of embodiment should evolve with technology. High embodiment and low embodiment are too vague to accurately reflect the current capabilities available to researchers and designers. Thus, the concept of high embodiment should be modernized to *human embodiment* to reflect that the agent possesses all the capabilities found in humans, such as facial expressions, eye gaze, lip synchronization, and body sway; and since all people use multiple gestures, the agent performs all gesture types. Human embodiment would suggest all human capabilities are present in the agent design. Designs that lack some aspects of human capabilities could be labeled *embodiment* to signal that the agent has some human-like features and capacity, but lacks the presentation of the full range of human gesturing abilities. Embodiment would require researchers to detail which aspects of the design are not representative of human potential. For example, an agent performing only deictic and beat gestures with no facial expressions would be considered embodied because it lacks other gesture types and shows no emotions commonly presented by humans. For agents that only use lip synchronization with no other capabilities, they would be considered *static* as they are currently categorized in the literature. Finally, agents that possess no elements of embodiment would be considered *pictures*.

However, research into human embodiment needs to be conducted to understand what design features need to be adjusted due to context. While there is evidence that enhanced gesturing is beneficial for foreign language users, native speakers may find enhanced gesturing a distraction because they do not need extra assistance with comprehension. However, complicated materials or processes might require agents to gesture more with native-speaking participants. These aspects, as well as the learning of declarative and procedural knowledge, need to be examined further in future research.

### Limitations

There are some limitations to this study. First, a majority of the students would be considered advanced foreign-language users of English. Therefore, further studies should examine whether other language levels would benefit from enhancing the gesture rate. It is possible that listening comprehension for students of higher ability would allow them to process enough of the spoken language to attend to the enhanced gesture frequency and benefit without being mentally overloaded. However, participants with lower proficiency in listening comprehension might not benefit from the enhanced gesture frequency because of the mental strain caused by trying to process the amount of auditory

and visual information presented. Therefore, these results may not be applicable to learners of all foreign language proficiencies.

In addition, this video did not include any visual aids or verbal redundancy that are used in most multimedia learning environments. It is unknown whether gestures would offer the same benefits with learning strategies present. Therefore, future research should examine whether enhancing the gesture rate with visual aids and verbal redundancy provides similar benefits for foreign language users.

## Conclusion

This research sought to examine how gesture frequency impacted agent perception and learning outcomes. This study provides more evidence that gesture types and frequency perform a critical role in agent perception and learning with foreign-language users. Similar to human-to-human gesture studies that found gesturing significantly helped foreign-language users (Church et al. 2004; Kelly et al. 2009), and that enhancing the gesture rate was more beneficial for students learning complex information (Alibali et al. 2001), the addition of all gesture types and frequency were essential in this study. The significance of both gesture types with the perception of agent persona shows that gestures have the potential to prime the social relationship between agents and humans. Likewise, enhancing the gesture rate might have more benefits when it comes to learning outcomes. This might result from gestures representing pictorial representations for the participant to “see” when listening to a presentation of declarative information without verbal redundancy or visual aids.

However, if embodiment does prime the social relationship and promote better learning outcomes, then agents should be designed to function similarly to humans in multimedia environments. This form of non-verbal communication could benefit other computer-based environments, such as virtual environments (VR), that rely on immersive capabilities to allow users to have realistic experiences. Thus, agents designed to replicate human gesture capabilities and frequencies could be fundamental for educators and designers creating multimedia and immersive experiences with agents that mimic real-life interactions with humans.

## Acknowledgements

Not applicable.

## Authors' contributions

RD conceptualize the study, built the intervention, collected data, and contributed to writing the manuscript. LW performed the data analysis and contributed to writing the manuscript. JV collected data and contributed to writing the manuscript. All authors read and approved the final manuscript.

## Funding

The authors did not receive funding for this manuscript.

## Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declaration

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Department of English Linguistics and Language Technology, Hankuk University of Foreign Studies, 107 Imun-ro, Imun-dong, Dongdaemun-gu, Seoul, South Korea. <sup>2</sup> English for International Conferences and Communication, Hankuk

University of Foreign Studies, 107 Imun-ro, Imun-dong, Dongdaemun-gu, Seoul, South Korea. <sup>3</sup> Business School, Information System and Operation Management, Hankuk University of Foreign Studies, 107 Imun-ro, Imun-dong, Dongdaemun-gu, Seoul, South Korea.

Received: 2 December 2020 Accepted: 15 March 2021

Published online: 20 April 2021

## References

- Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: a meta-analysis. *Journal of Educational Psychology, 104*(1), 250.
- Alibali, M. W., Heath, D. C., & Myers, H. J. (2001). Effects of visibility between speaker and listener on gesture production: some gestures are meant to be seen. *Journal of Memory and Language, 44*(2), 169–188.
- Brown, J. L. (1995). *Observing dimensions of learning in classrooms and schools*. ASCD.
- Carlotta, T., & Jaques, P. A. (2016). The effects of animated pedagogical agents in an English-as-a-foreign language learning environment. *International Journal of Human-Computer Studies, 95*, 15–26.
- Carrier, K. (1999). The social environment of second language listening: does status play a role in comprehension? *The Modern Language Journal, 83*, 65–79.
- Choi, S., & Clark, R. E. (2006). Cognitive and affective benefits of an animated pedagogical agent for learning English as a second language. *Journal of Educational Computing Research, 34*(4), 441–466.
- Church, R. B., Ayman-Nolley, S., & Mahootian, S. (2004). The role of gesture in bilingual education: does gesture enhance learning? *International Journal of Bilingual Education and Bilingualism, 7*(4), 303–319.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46.
- Davis, R., & Antonenko, P. (2017). Effects of pedagogical agent gestures on social acceptance and learning: virtual real relationships in an elementary foreign language classroom. *Journal of Interactive Learning Research, 28*(4), 459–480.
- Davis, R. O. (2018). The impact of pedagogical agent gesturing in multimedia learning environments: a meta-analysis. *Educational Research Review, 24*, 193–209.
- Davis, R. O., & Vincent, J. (2019). Sometimes more is better: agent gestures, procedural knowledge and the foreign language learner. *British Journal of Educational Technology, 50*(6), 3252–3263.
- Fountoukidou, S., Ham, J., Matzat, U., & Midden, C. (2019). Effects of an artificial agent as a behavioral model on motivational and learning outcomes. *Computers in Human Behavior, 97*, 84–93.
- Grèzes, J., Costes, N., & Decety, J. (1999). The effects of learning and intention on the neural network involved in the perception of meaningless actions. *Brain, 122*(10), 1875–1887.
- Griffiths, R. (1990). Facilitating listening comprehension through rate-control. *RELC Journal, 21*(1), 55–65.
- Hostetter, A. B. (2011). When do gestures communicate? a meta-analysis. *Psychological Bulletin, 137*(2), 297–315.
- Hostetter, A. B., & Skirving, C. J. (2011). The effect of visual vs. verbal stimuli on gesture production. *Journal of Nonverbal Behavior, 35*(3), 205–223.
- Johnson-Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., & Grafton, S. T. (2003). Actions or hand-object interactions? Human inferior frontal cortex and action observation. *Neuron, 39*(6), 1053–1058.
- Kelly, S. D., McDevitt, T., & Esch, M. (2009). Brief training with co-speech gesture lends a hand to word learning in a foreign language. *Language and Cognitive Processes, 24*(2), 313–334.
- Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge: Cambridge University Press.
- Krauss, R. M., Morrel Samuels, P., & Colasante, C. (1991). Do conversational hand gestures communicate? *Journal of Personality and Social Psychology, 61*(5), 743–754.
- Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston: Heinle & Heinle Pub.
- Marzano, R. J., & Pickering, D. J. (1997). *Dimensions of learning: Teacher's manual*. Assn for Supervision & Curriculum.
- Mayer, R. E. (2014). *Based principles for designing multimedia instruction. Acknowledgments and Dedication, 59*.
- Mayer, R. E., & DaPra, C. S. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied, 18*(3), 239.
- Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: role of speaker's voice. *Journal of Educational Psychology, 95*(2), 419–425.
- McCafferty, S. G. (2002). Gesture and creating zones of proximal development for second language learning. *The Modern Language Journal, 86*(2), 192–203.
- McNeill, D. (1992). *Hand and mind: what gestures reveal about thought*. Chicago: University of Chicago Press.
- Reeves, B., & Nass, C. (1996). *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge: Cambridge University Press.
- Ritzhaupt, A. D., & Barron, A. (2008). Effects of time-compressed narration and representational adjunct images on cued-recall, content recognition, and learner satisfaction. *Journal of Educational Computing Research, 39*(2), 161–184.
- Ritzhaupt, A. D., Barron, A. E., & Kealy, W. A. (2011). Conjoint processing of time-compressed narration in multimedia instruction: the effects on recall, but not recognition. *Journal of Educational Computing Research, 44*(2), 203–217.
- Ritzhaupt, A. D., & Kealy, W. A. (2015). On the utility of pictorial feedback in computer-based learning environments. *Computers in Human Behavior, 48*, 525–534.
- Schlaug, G., Knorr, U., & Seitz, R. J. (1994). Inter-subject variability of cerebral activations in acquiring a motor skill: a study with positron emission tomography. *Experimental Brain Research, 98*(3), 523–534.
- Schroeder, N. L., & Adesope, O. O. (2013). How does a contextually-relevant peer pedagogical agent in a learner-attenuated system-paced learning environment affect cognitive and affective outcomes? *Journal of Teaching and Learning with Technology, 2*(2), 114–133.
- Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How effective are pedagogical agents for learning? A meta-analytic review. *Journal of Educational Computing Research, 49*(1), 1–39.

- Schroeder, N. L., Yang, F., Banerjee, T., Romine, W. L., & Craig, S. D. (2018). The influence of learners' perceptions of virtual humans on learning transfer. *Computers & Education*, *126*, 170–182.
- Theune, M., & Brandhorst, C. J. (2009). To beat or not to beat: beat gestures in direction giving. In S. Kopp & I. Wachsmuth (Eds.), *Gesture in embodied communication and human-computer interaction* (pp. 195–206). Springer.
- Ullman, M. T. (2016). The declarative/procedural model: a neurobiological model of language learning, knowledge, and use. In *Neurobiology of language* (pp. 953–968). Academic Press: Cambridge.
- Wang, J., Dawson, K., Saunders, K., Ritzhaupt, A. D., Antonenko, P. P., Lombardino, L., Beal, C., Keil, A., Dogan, N., Luo, W., Cheng, L., & Davis, R. O. (2018). Investigating the effects of modality and multimedia on the learning performance of college students with dyslexia. *Journal of Special Education Technology*, *33*(3), 182–193.
- Wang, F., Li, W., Xie, H., & Liu, H. (2017). Is pedagogical agent in multimedia learning good for learning? A meta-analysis. *Advances in Psychological Science*, *25*(1), 12–28.
- Willems, R. M., & Hagoort, P. (2007). Neural evidence for the interplay between language, gesture, and action: a review. *Brain and Language*, *101*(3), 278–289.
- Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(6), 1047–1060.
- Woo, H. L. (2008). Designing multimedia learning environments using animated pedagogical agents: factors and issues. *Journal of Computer Assisted Learning*, *25*(3), 203–218.
- Yung, H. I., & Paas, F. (2015). Effects of cueing by a pedagogical agent in an instructional animation: a cognitive load approach. *Journal of Educational Technology & Society*, *18*(3), 153–160.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---