



Computer-based Human Agent Image and Modality Effects – Does it Facilitate Deeper Information Processing?

Li Xuanxi, Fong Soon Fook, *Balakrishnan Muniandy, **Hanafi Atan & **Rozhan M. Idrus

School of Educational Studies, Universiti Sains Malaysia, 11800 Penang, MALAYSIA

**Centre for Instructional Technology and Multimedia, Universiti Sains Malaysia, 11800 Penang, MALAYSIA*

***School of Distance Education, Universiti Sains Malaysia, 11800 Penang, MALAYSIA*

Abstract

This study investigated the image and modality effects of a narrating human agent on the learning of Chinese idioms among Malaysian primary school students. A total of 138 students were randomly assigned to one of the three modes of instruction. The results find support for the image effect and the modality effect on the learning of Chinese idioms, and suggest that incorporating a narrating human agent with facial expressions, gestures and intonation can enhance deeper information processing.

Xuanxi, L., Fong, S.F., Muniandy, B., Atan, H. & Idrus, R.M. (2010). Computer-based Human Agent Image and Modality Effects – Does it Facilitate Deeper Information Processing? *Malaysian Journal of Educational Technology*, 10(2), pp. 11-18.

Introduction

As communication between China and the world is becoming increasingly open and frequent, learning Chinese language has become more and more important. Chinese idioms have been and are still the core aspect of Chinese language today. As Luk & Ng (1998) stated, "Idioms derived from allusions, poetry and proverbs have cultural, literary, moral and social connotations that maintain a sense of cultural identity amongst Chinese people and create a sense of continuity to their history".

Chinese idioms are difficult to understand in depth for primary learners. Xinfang (2005) indicated that primary L1 learners (learners who learn Chinese language as their mother tongue) have problem to understand the profound meanings of idioms thoroughly. For instance, Luk & Ng (1998) found that primary students in Hong Kong seldom use Chinese idioms and always use them incorrectly and inappropriately. In Malaysia, a preliminary survey was conducted on the teaching of Chinese idioms among 36 Chinese language teachers from 14 Chinese primary schools indicated that the learning of Chinese idioms among primary students in Malaysia was also a problem area.. These teachers indicated that just only 49.03% students could understand and apply idioms correctly.

According to Mayer and Clark (2003), pedagogical agents can be representations of real people using video and human voice or artificial characters using animation and computer-generated voice. Xuanxi et al (2007) explained that it is difficult for artificial pedagogical agent to express messages in language learning through dialogues alone but it could be effectively conveyed through a human agent that have corresponding facial expressions and gestures. Lester (2000) also indicated that integrating a human-like pedagogical agent into a learning environment can provide four important educational benefits – may encourage the learners to care about their own progress; may convey enthusiasm for the learning of the subject matter at hand and may foster similar levels of enthusiasm in the learner; may make learning fun



and may consequently encourage them to spend more time in the learning environment. Ying (2006) investigated the effect of a human agent in an e-learning program on the pronunciation of Chinese language entitled "Er Hua". She indicated that the computer-based human agent have a significant effect on L2 learners (learners who learn Chinese language as their second language). Xuanxi et al (2007) explained that it is difficult for animated pedagogical agent to express messages in language learning through dialogues alone but it could be effectively conveyed through a human agent that have corresponding facial expressions and gestures. Mayer and Clark (2003) stated that pedagogical agents can be representations of real people using video and human voice or artificial characters using animation and computer-generated voice. It is suggested that a human-agent integrated into the learning process can help students understand the profound meanings of Chinese idioms better. In this study, the human agent is defined as a talking-head video with facial expressions, gestures and intonations.

Overview of Present Study

The current study was designed to examine the image effect and the modality effect on the learning of Chinese idioms among Malaysian primary school students. Two questions were addressed:

- Was there an image effect?
- Was there a modality effect?

The image effect suggests that students presented with the narration + human agent condition will outperform their peers presented with the narration condition. The modality effect suggests that students using the narration condition will benefit more from the increased working memory and more likely to remember and to use the instruction content of the lesson (Dunsworth & Atkinson, 2007; Atkinson, 2002).

There are several reasons for integrating a computer-based human agent into the learning of Chinese idioms in this study. Firstly, the theoretical perspectives that help account for the effect of human agent are the social agency theory (Atkinson et al, 2005; Moreno et al, 2001) and Mayer's dual-channel information processing theory (Mayer, 2001). The Social agency theory contends that with the integration of a narrating human agent and non-verbal social cues (e.g., gestures, gaze, and facial expression) into an e-learning program could simulate the human-to-human connection, therefore facilitating the students' engagement in the learning process. The dual-channel information processing in Mayer's cognitive theory of multimedia learning (2001) suggests that we can process visual and aural information through two channels simultaneously – the verbal channel for processing words and the visual channel for processing pictures. Words can be presented either aurally (i.e., narration) or visually (i.e., on-screen text); but images can only be presented in the visual version, thus, when both words and images are involved in an instruction, verbal information presented aurally rather than visually can decrease the cognitive load and increase the size of working memory. Secondly, based on Piaget's (1951) theory, the subjects of this study belong to the third stage of cognitive development – the concrete operational stage (7-11 years old). During this stage, children are capable of various logical operations but only with concrete things. They are able to understand complex logical ideas - classification, grouping and ordering problems. However, the limitation of this stage is that the children need concrete representations to tie to their thinking; they are not yet fully aware of the principles involved. Children at the age group of 7-11 years old still have problems with highly abstract thought (Liangfang, 2001; Gage & Berliner, 1998). It is suggested that a computer-based human agent with social cues (speech, gestures, and facial expressions) can enhance children's deeper processing and hence learning on Chinese idioms.

The independent variables of the present study are the three modes of e-learning programs: On-screen Text + Graphics + Captions (TGC) mode, Narration+ Graphics + Captions (NGC) mode, and Human Agent + Narration + Graphics + Captions (HANGC) mode. The human agent's facial expressions and gestures were coordinated with the monologue narration and were designed to tell the story of each idiom. For the dependent variables, the posttest scores and the scores of the attitude scores were manipulated to measure the performance and the level of software preferences. The post scores were selected as an indicator of academic performance for all groups participated in the study. The level of programme preferences of using the three modes was measured using the attitude test.

Research Hypothesis

H1: Students using HA+narration+graphics+captions (HANGC) mode will attain significantly higher post scores (PS) than students using narration+graphics+captions (NGC) mode, that is, PSHANGC > PSNGC

H2: Students using narration+graphics+captions (NGC) mode will attain significantly higher post scores (PS) than students using on-screen text +graphics+captions (TGC) mode, that is, PSNGC > PSTGC

There were 138 subjects from two Chinese primary schools in the city of Penang in Malaysia. All of the subjects were taken from their normal intact classes. Their ages ranged from 10-11 years old. The participants were randomly assigned to one of the three e-learning programs: TGC, NGC or HANGC modes.

E-Learning Interface

Each program included the same 10 Chinese idioms which had been chosen from the primary four Chinese idioms syllabus. Every program had two units, and every unit had 5 Chinese idioms. The e-learning program could be divided into three sections (Fig. 1): Section A is a list of idioms; Section B is the basic knowledge of each idiom such as definition, derivation, synonyms, antonyms, and example usage; Section C is the story of each idiom.

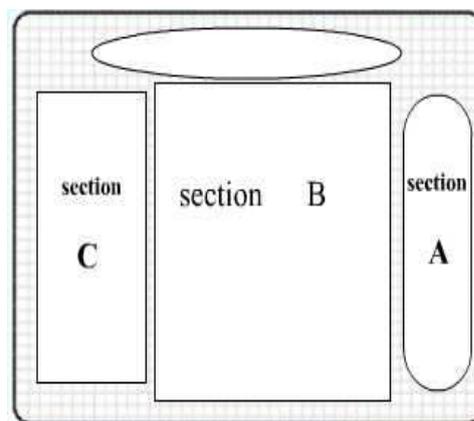


Figure 1 Three sections of the e-learning program in this study

The three programs which were on-screen text mode (TGC), narration mode (NGC), and human agent + narration mode (HANGC) differed from one another only in the versions of presentations on stories of idioms in section C. In the TGC mode, the learners were presented with the story of each idiom in text

version in section C (Fig. 2). In the NGC and HANGC modes, the same section A and section B were presented in Graphics and Captions (GC) but the idiom stories in section C were presented respectively through a monologue-style human voice and through a human agent plus a monologue-style human voice. The script used to create the voice files for the NGC and HANGC conditions was identical to the text in the NGC versions. Presented with the NGC mode, the students listened to the audio instruction and watched the Graphics and Captions (GC) on the screen which was synchronized to the voice (Fig. 3). Presented with the HANGC mode, the students listened to the audio instruction which was expressed by the HA and watched her gestures and facial expressions which were synchronized to the GC. The HA was the video of a teacher that uses facial expressions and gestures to direct learner attention as she talks (Fig. 4).



Figure 2 On-screen text + Graphics +Captions condition (TGC)



Figure 3 Narration Condition + Graphics + Captions condition (NGC)



Figure 4 Narration + Human Agent+ Graphics +Captions condition (HANGC)

Instruments

Two main instruments were used to collect data for this study. One instrument was Chinese idioms performance test (pretest and posttest) to evaluate the students' Chinese idioms performance. The other instrument was an attitude test used to assess student's perception of the three modes of e-learning programmes. The test items in the posttest were the same as that of the pretest except that the sequence of questions and answers were altered so as to reduce 'item memory practice'. An item analysis was carried out based on the results of the testing and employed two types of information – item difficulties, and discrimination indices – to improve the test. Content validity of the 25 items test was established by 3 Chinese language specialists who designed the instrument, 20 items were selected at last. Reliability of the test was estimated by using the Cronbach Alpha procedure.

Results

T-test and Analyses of Covariance (ANCOVA) were used in analyzing the data. Table 1 shows a difference between the means of post scores for group with TGC and group with the NGC, the mean post scores for group with the NGC (65.00) was higher than the mean gain score for group with the TGC (55.43), and $p=0.041$ which showed that the difference was significant. The mean post scores for group with the HANGC (77.07) was higher than the mean gain score for group with the NGC (65.00), and $p=0.001$ which indicated that the difference was significant.

Table 1 T-test of the posttest scores of students in the various treatment groups

	MODE	N	Mean	Std. Deviation	df	t	Sig.
Posttest Scores	TGC	46	55.43	24.627	90	-2.07	0.041
	NGC	46	65.00	19.379		0	
	NGC	46	65.00	19.379	90	3.498	0.001
	HANGC	46	77.07	13.105			

Significance: * $P < 0.05$

In order to remove the effect of pretest, the pretest was used as the covariate. ANCOVA was used to determine if the differences which were found in the mean scores of post-test were significant or not. Table 2 indicates the results of test of statistical significance on the differences observed in the mean scores of the posttest for the various treatment groups with $F=31.391$, Mean Square=5337.386, and $p=0.000$. Therefore, these differences in the post scores were significant.

Table 2 ANCOVA of the posttest scores of students in the various treatment groups

Dependent Variable: Posttest scores

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	39945.070(a)	3	13315.023	78.310	.000



Intercept	32840.996	1	32840.996	193.148	.000
PRE	29136.012	1	29136.012	171.357	.000
MODE	10674.772	2	5337.386	31.391	.000
Error	22784.097	134	170.031		
Total	660825.000	138			
Corrected Total	62729.167	137			

a R Squared = .637 (Adjusted R Squared = .629)

The results are summarized as follows:

- Post scores (PS) attained by students using the HA +Narration +Graphics +Captions (HANGC) mode was significantly higher than that attained by students using Narration + Graphics + Captions (NGC) mode, that is, PSHANGC >PSNGC.
- Post scores (PS) attained by students using the Narration + Graphics + Captions (NGC) mode was significantly higher than that attained by students using the On-screen text + Graphics + Captions (TGC) mode, that is, PSHANGC >PSNGC
- There is no significant difference between the attitude scores (AS) attained by students using the HA + Narration + Graphics + Captions (HANGC) mode and students using the Narration + Graphics + Captions (NGC) mode.
- There is no significant difference between the attitude scores (AS) attained by students using the Narration + Graphics + Captions (NGC) mode and students using the On-screen text + Graphics + Captions (TGC) mode.

Discussion

Is there an image effect? Hypothesis 1 was supported. This study revealed that a human agent's visual presence in sharing the stories of idioms aurally (HANGC) could foster learning. This result is supported by other research studies on animated pedagogical agent (Dunsworth & Atkinson, 2007; Atkinson, 2002) which indicated that learners presented with the image of a pedagogical agent outperformed their peers on performance who were not presented with the image (Moreno, 2001).

There are several possible reasons for this result: Firstly, the possible reason for the positive image effect on students' performance could be based on Mayer's cognitive theory of multimedia learning (2001). According to the dual-channel information processing, when stories of idioms were presented by both the human agent with narration, the visual presence of human agent was processed through the visual channel while the narration information was processed through the verbal channel. There was an increased working memory because the cognitive architecture received information from both auditory and visual channels simultaneously (Dunsworth & Atkinson, 2007; Atkinson, 2002). According to Piaget (1951)'s cognitive stage, the intellectual development process of the subjects with the age 9-10 in this study belong to the concrete operational stage (ages 7 to 11). Children still have problems with highly abstract thought and they need concrete representations to tie to their thinking. As a result, children may not thoroughly understand the profound meanings of idioms just by the on-screen text presentation. By using social cues such as gestures and facial expressions, the human agent could direct the children's attention and help them to combine verbal and non-verbal forms of communication (Dunsworth & Atkinson, 2007). Thirdly, another possible reason could be based on the characteristics of Chinese idioms such as regular



construction, complex contents, and profound meanings which make idioms difficult to understand for students. Xuanxi et al (2007) indicated “for some subjects such as language, the image of an agent plays a fundamental role during the learning process. Facial expressions can add vital information about the intensity and valence of the social events described”. When a human agent with gestures and facial expressions was integrated into the learning process to explain the story of each idiom, the sophisticated messages relayed by the idioms were interpreted by the human agent, and then students could understand the idioms deeply.

Was there a modality effect? Hypothesis 2 together with Hypothesis 4 in this research predicted that instruction containing aural explanations for the stories of idioms had significantly positive effects on both students’ performance and attitudes than instruction with textual explanations (modality effect). Hypothesis 2 was supported. This study revealed that the e-learning program containing narrated stories of idioms (NGC) was more effective at promoting learning than the program with textual stories of idioms (TGC). This result is also supported by other researches on animated pedagogical agent (Atkinson, 2002; Moreno, 2001) which indicated that students learned better from visual and verbal presentations when the verbal information was presented as speech rather than as on-screen text.

Was there a modality effect? Hypothesis 2 was supported. It revealed that the e-learning program containing aural stories of idioms (NGC) was more effective at promoting learning than the program with textual stories of idioms (TGC). This result is also supported by other researches on animated pedagogical agent (Atkinson, 2002; Moreno, 2001) which indicated that students learned better from visual and verbal presentations when the verbal information was presented as speech rather than as on-screen text.

There are several reasons for this result: Firstly, the possible reason for the significantly positive modality effect on students’ performance can be explained by Mayer’s (2001) cognitive theory of multimedia learning. According to the design of the program NGC and TGC, there were the same 10 idioms embedded in each instruction, of which long stories were presented respectively in narration version and in visual version (on-screen text). Each idiom was presented individually, and all the basic knowledge of idiom including the definition, derivation was presented in visual version (on-screen text), when the long story of the idiom was also presented in visual version, because of the limited capacity of the visual channel, too much visual information may hamper the learning of knowledge. Whereas when the long story of the idiom was presented in aural version, according to the dual-channel information processing theory, the working memory capacity was increased by both verbal and visual channels, and then students get better learning result. Furthermore, every story of idioms was independent, did not relate to other story, and it was not necessary for students to remember each word of the story. According to Xuanxi et al (2007), when the auditory information is not complex and not correlated with each other, and the learning does not constantly require reference to the aural information, the advantage of increased working memory capacity derived from dual-channels information processing may be apparent.

Conclusion

This study suggested that incorporating a computer-based human agent could play a very important communication role in virtual learning environments. As an engaging character, the human agent with facial expressions, gestures and intonations could help focus students’ attention in relation to deeper cognitive learning process. This study also found support for the modality effect in the process of Chinese



idioms learning, that is, students learn better when the story of idioms presented in narration version rather than on-screen text version.

References

- Atkinson, R. (2002). Optimizing Learning From Examples Using Animated Pedagogical Agents. *Journal of Educational Psychology, 94*, pp. 416-427.
- Atkinsin, R. K., Mayer, R. & Merrill, M. (2005). Fostering Social Agency In Multimedia Learning: Examining The Impact Of An Animated Agent's Voice. *Contemporary Educational Psychology, 30*, pp. 117-139
- Bloom, B. S. (1984). The 2 Sigma Problem: The Search For Methods Of Group Instruction as Effective As One-To-One Tutoring. *Educational Researcher, 13*, pp. 4-16.
- Clark, C. R. & Mayer, R. (2003). *E-Learning and the Science Of Instruction: Proven Guidelines For Consumers And Designers Of Multimedia Learning*. San Francisco, Pfeiffer.
- Dunsworth. Q. & Atkinson, R. K. (2007). Fostering Multimedia Learning Of Science: Exploring The Role Of An Animated Agent's Image. *Computers and Education, 49*(3), pp. 677-690.
- Gage, N. L. & David C. Berliner (1998). *Educational Psychology*. Boston: Houghton Mifflin Company.
- Liangfang, S. (2001). *Learning Theory*. Beijing, Public Education
- Luk, R. W. P., & Ng, A. B. Y. (1998). Computer - Assisted Learning Of Chinese Idioms, *Journal of Computer Assisted Learning, 14*, 2-18.
- Mayer, R. E. (2001). *Multimedia Learning*. New York, Cambridge University Press.
- Moreno, R. (1999). *Introducing Social Cues In Multimedia Learning: The Role Of Pedagogical Agents' Image And Language In A Scientific Lesson*. Santa Barbara, University of California. Unpublished Doctoral Dissertation.
- Moreno, R. (2001). Cognitive And Motivational Consequences Of Adapting An Agent Metaphor In Multimedia Learning. Do The Benefits Outweigh The Costs? *Proceedings of WebNet 2001*, (pp. 873-878). Charlottesville, VA: Association for the Advancement of Computers in Education.
- Moreno, R., Mayer, R.E., Spires, H., & Lester, J. (2001). The Case For Social Agency In Computer-Based Teaching: Do Students Learn More Deeply When They Interact With Animated Pedagogical Agents? *Cognition and Instruction, 19*, pp. 177-213.
- Peck, D. (1998). "Teaching Culture: Beyond Language." Retrieved Jan. 20, 2008, from <http://www.yale.edu/ynhti/curriculum/units/1984/3/84.03.06.x.html>
- Piaget, J. (1951). *The Child's Conception Of Physical Causality* (M. Gabain, Trans.). New York, Humanities Press. (Original work published 1930.).
- Slater, D. (2000). *Interactive Animated Pedagogical Agents: An Introduction To An Emerging Field*. Stanford University: <http://ldt.stanford.edu/~slater/pages/agents/main.htm>
- Xinfang, J. (2005). Four Problems In The Teaching Of Chinese Idioms: <http://www.yxedu.com/html/2005-12-1/2005121134455.htm> (Retrieved January 26, 2006).
- Xuanxi, L., Fong Soon Fook & Rozhan M. Idrus (2007). Delivery Modalities Of Animated Pedagogical Agents In E-Learning: A Review Of Empirical Researches. *Malaysian Journal of Educational Technology, 7*(2), pp. 65-74.
- Ying, L. C. (2006). *The Effect Of Pedagogical Agent On The Mean Score Of Posttest Of Students With Different Achievement Level In The Learning Of "Er Hua"*. Master Thesis (unpublished), USM