別紙１

論文内容の要旨

論文題目:

Reactive Animated Pedagogical Agents:　Exploring Dyadic Gaze Interaction
(反応的ペダゴジカル・エージェント：視線インタラクションの研究)

氏名　　李 翰 柱 （ Ｌｅｅ Ｈａｎｊｕ）

The main goal of this dissertation is to implement reciprocal features of social interactions with animated pedagogical agents and verify their learning effects. Pedagogical agents are artificial entities designed to facilitate learning, often taking form as characters in computer screen. Pedagogical agent was announced as a potential tool to innovate learning, by incorporating social aspect of learning environment by utilizing social cues, such as gestures, voice and gaze. The dissertation covers two major reciprocal social features, temporal contingency and joint attention.

The dissertation consists of nine chapters. In chapter 1, a brief review of current literature and theoretical basis of this project is provided. Four major theories are introduced. Firstly, the persona effect theory claims that the mere physical presence of the pedagogical agent leads to better learning. The theory is based on a study which is criticized for not comprising a control group, but is still frequently cited among the literature. The theory is not yet entirely rejected or confirmed, as conflicting evidences surrounding the persona effect have been found. Secondly, the split-attention effect theory claims the opposite of the persona effect theory, stating that pedagogical agents may split the limited attention of learners thus causing harm on learning. Thirdly, the modality effect

theory states that by simultaneously using dual channel—visual, audio—inherent in human working memory, the learner can process more information in limited time. Finally, the social agency theory claims that incorporating social agency in computer-based learning facilitates deeper learning from learners by eliciting social stance. Pedagogical agent, anthropomorphic image of computer system, is largely supported by the social agency theory. As result, pedagogical agent literature has focused on making more human-like agents, in attempt to elicit more social agency. Nonverbal social cues such as gesture and gaze have also received attention from literature. Several studies reported positive effect of social cues. However, the quality assessed by previous studies is limited to the life-likeness of agents' animation. That is, while human nonverbal social cues are reciprocal, no research has been made on how pedagogical agents should react to learners' social cues.

Despite of high claims made by initial studies, the pedagogical agent research has not yet succeeded to yield decisive results, partially due to the fact that some of the crucial features of human social communication remain neglected. To fill in the gap, the dissertation focuses on reciprocity, which is one of major characteristics of human social communication. Recent developments in social neuroscience suggest that social cognition may be fundamentally different when individuals are interacting with others rather than merely observing. Moreover, reciprocal interaction is argued to be strongly related to human learning, as studies from the field of developmental psychology report that live social interaction facilitates learning of young children, even when the interaction is provided through screen media.

In chapter 2, I explain the design and implementation of the experimental-purpose pedagogical agent used in the project. The agent designed to be easily modifiable, easily extendable, while maintaining strict control over the behaviours of the agent itself. Using the agent, I conducted five experiments detailed in the following chapters.

In chapter 3, the persona effect theory and the split attention effect theory is tested (experiment 1). Research surrounding two theories has produced mixed results. This may have been affected by the great variety in functions and properties of agents used in the previous studies. Using an agent with only the most basic functions, I demonstrated that mere presence of pedagogical agent is not enough to make a difference in learning.

In chapter 4, temporal contingency is implemented into gaze interaction of the pedagogical agent, and its learning effect is tested. The agent simulated mutual gaze, gaze following, and joint attention with students while teaching foreign language words. The result demonstrated that temporally contingent gaze interaction improves the learning with animated pedagogical agents. The eye tracking data analysis revealed that a certain time-window of fixation duration on learning materials and the agent was related to higher test scores. The average fixation duration of the higher scoring group (split by the test score median) was inside 350~750ms. Based on the result, I propose two hypotheses in attempt to explain the result. The first is that lack of temporal contingency

induced larger cognitive load by inflicting the need for more frequent visual search. The seconds is that the temporal contingency of gaze interaction may have yielded greater social presence, triggering different process while learning.

To assess further, I conducted experiment using non-anthropomorphic agents, to test if temporal contingency affects non-social cues. In chapter 5, I tested if an arrow shaped agent produce similar results as the anthropomorphic agent (experiment 3). The result revealed that temporal contingency did not affect learning from the arrow agent. To confirm that this was not caused by the difference of saliency between the agents, I conducted experiment 4 (explained in chapter 6) using a smaller arrow agent with lesser saliency. The learning effect of the agent was also not affected by temporal contingency, indicating that saliency was not the cause of the result. This implies that temporal contingency effect is unique to social cues. Our hypotheses on the reason of temporal contingency effect, which were raised in chapter 4 was: 1) temporal contingency reduces extraneous cognitive load related to visual search, 2) temporal contingency prime social stance in learners which enhance learning. The basis for hypothesis 1 was that average fixation duration was shorter when the interaction was not temporally contingent. This relation held true for all three agent types, but for the arrow agents, it was not correlated with learning outcomes. Also, when agents were not contingent, average fixation duration of anthropomorphic agent group was the longest, but word learning score was the lowest. This suggests that cognitive load caused by visual search was unlikely the cause of temporal contingency effect.

The post-hoc analysis showed that agent type had no effect on learning outcomes when the agents were temporally contingent. However, when they were not temporally contingent, the anthropomorphic agent yielded less learning effect compared to non-human agents. I argue that this may have been caused by incomplete activation of gaze-related social attention system. When social cues are not directed at one self, it is no longer reciprocal and it is natural for the person to neglect the social cues. Thus, when the anthropomorphic agent displayed non-reciprocal (non-temporally contingent) gaze, the natural neurocognitive function of participants was to neglect the gaze. On the other hand, non-social cues are usually not reciprocal and adults are trained to use non-social cues for learning. In sum, participants had to provide extra effort when the anthropomorphic agent displayed non-reciprocal (non-temporally contingent) gaze, as it was against their natural neurocognitive system for social cues, whereas non-social cues did not suffer as they used different system.

To assess deeper into temporal aspect of reciprocal interaction, I discuss how temporal order of contribution in joint attention may affect learning in chapter 7. Previous studies from developmental science have reported that whether the referenced object was already the focus of attention of the child, when forming joint attention had significant effect on child language learning. This was labeled maternal ‘follow-in’ vs ‘lead-in’; follow-in refers to following the child’s attention,

whereas lead-in refers to the mother leading the child's attention. However, adults are much more capable of redirecting their attention compared to children. Thus I conducted experiment 5 to assess whether 'follow-in' vs 'lead-in' affects adult learning. The experiment 5 used an anthropomorphic agent that could either lead or follow learner's gaze to form joint attention. The result indicates that 'follow-in' vs 'lead-in' does not affect adult learning in a major way.

In chapter 8, I present general discussion of this project. The five experiments introduced in this dissertation led to conclude that 1) The mere physical image of pedagogical agent did not facilitate learning, 2) For social cues, temporal contingency facilitated learning, 3) For non-social cues, temporal contingency did not facilitate learning, 4) Social cues were not necessarily superior compared to non-social cues, and 5) The short-term learning effect of temporal order of contribution in joint attention may be small for adults. I provide a possible explanation on why the temporal contingency effect is specific to social-cues; as non-social cues are generally not reciprocal, adults are trained to react to non-social cues that are not-contingent. By comparison, non-contingent social interaction may trigger incomplete activation of social attention system, thus demanding more effort from recipients.

On conclusion 4) social cues were not necessarily superior compared to non-social cues, I argue that this may be a result of training, or adaptation. This is supported by the fact that young children are affected more by the lack of social cues. Indeed, children under three years of age learn less from screen media than from engaging in live social interaction with adults. Moreover, the effect is mitigated when live social interaction is provided through screen media. While two types of cues used in this study (arrow and gaze) are known to elicit similar response, attention systems behind each cue are argued to be different. Moreover, several studies indicate that human infants are born with ability to utilize social cues to gather meta-information crucial for learning, as seen from their specific response to social cues. Indeed, humans may have dedicated or unique attention system for social cues, which is strongly connected to learning from early stages.

In chapter 9, I discuss the limitation and future works of this project and conclude this dissertation.

This study provides first empirical evidence that reciprocal aspect of social interaction has practical implications for computer-based education. Also, the findings from this project may contribute to multiple academic fields, cognitive science, cognitive neuroscience, and education engineering.