

**DESIGNING CHATBOT INTERFACES FOR LANGUAGE  
LEARNING: ETHNOGRAPHIC RESEARCH INTO AFFECT  
AND USERS' EXPERIENCES**

by

Yifei Wang

B.Eng., Beijing University of Chemical Technology, 1996  
M.A., East China University of Science and Technology, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Technology Studies Education)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2008

© Yifei Wang, 2008

## **Abstract**

During the past few decades, there has been increasing attention to multimodal adaptive language learning interface design. The purpose of this study was to examine users' experiences with a chatbot language learning interface through the lens of cognitive emotions and emotions in learning. A particular focus of this study was on users' interactions with a chatbot in a public setting and in a private environment. Focusing on the event of users' interaction with a chatbot interface, seventy-five interactions were videotaped in this study, in which fifteen users were asked to interact with the chatbot "Lucy" for their language learning. The video-stimulated post interaction interviews with participants provided complementary data for understanding their experiences with the language learning system. Analysis of twenty-five interactions selected from a total of seventy-five revealed five main factors of chatbot language tutor interface design and their relative significance in the process of users' meaning making and knowledge construction. Findings showed that users' sensory, emotional, cultural, linguistic and relational engagement influenced their responses to the chatbot interface, which in turn, shaped their learning processes. Building on a theoretical framework of cognitive emotions and emotions in learning, this study documented users' language learning processes with the chatbot language learning interface by investigating users' experiences. The findings and techniques resulting from this study will help designers and researchers achieve a better understanding of users' experiences with technology and the role of emotions in the processes of learning when using technology and assist them to improve the design of language learning environments.

# Table of Contents

<b>Abstract.....</b>	<b>ii</b>
<b>Table of Contents .....</b>	<b>iii</b>
<b>List of Tables .....</b>	<b>vi</b>
<b>Lists of Figures .....</b>	<b>vii</b>
<b>Preface.....</b>	<b>viii</b>
<b>Acknowledgements.....</b>	<b>x</b>
<b>Chapter One: Introduction .....</b>	<b>1</b>
1.1 Research Rationale.....	3
1.2 Research Purpose .....	5
1.3 Research Problems .....	7
1.4 Research Questions .....	9
1.5 Limitations of The Study.....	11
1.6 Thesis Structure.....	12
<b>Chapter Two: Literature Review.....</b>	<b>14</b>
2.1 The History of Computer – Assisted Language Learning .....	15
2.2 Chatbot Technology Overview .....	15
2.2.1 Chatbot technology development .....	15
2.2.2 Chatbot technology usage.....	18
2.3 Adaptivity.....	19
2.3.1 The ontology of emotion – Emotion research traditions .....	22
2.3.2 Evans’ accounts of emotion.....	27
2.3.3 State of the art of key research .....	31
2.3.4 State of the art of key projects.....	33
2.3.5 Skepticism towards Affective Computing.....	35
2.4 Multimodality.....	36
2.5 Emotions in Learning.....	37
2.5.1 The model of emotion in learning .....	40
2.5.2 Emotions in the service of cognition .....	41
2.6 Affective Domain of Learning .....	45
2.7 Conclusion .....	47
<b>Chapter Three: Research Design.....</b>	<b>48</b>
3.1 Methods.....	49
3.1.1 Pre-research Interview (Semi-structured).....	50
3.1.2 Participant Reconstructive Interviews .....	51
3.1.3 Participant Observation .....	52
3.1.4 Videotaping .....	53
3.2 Unit of Analysis: The Participants .....	54
3.3 Site of Research.....	55
3.4 Instrument for the Study.....	56
3.4.1 Chatbot Lucy .....	56
3.4.2 Cameras .....	59
3.4.3 Microphones.....	59
3.4.4 Tripod.....	59
3.4.5 Monitors .....	60
3.5 Participant Ethics .....	60
3.6 An Introduction to Data.....	61

3.6.1	Data collection.....	61
3.6.2	Data Analysis .....	64
3.7	Conclusion .....	65
<b>Chapter Four: Data Analysis and Findings .....</b>		<b>66</b>
4.1	A Narrative Approach to Analysis .....	66
4.2	Participants Selection in Narration.....	72
4.3	Lee’s Sensory Experiences.....	72
4.3.1	General Background of Lee.....	73
4.3.2	The Research Setting.....	74
4.3.3	Lee’s Sensory Experiences with Lucy.....	74
4.3.4	The Sensory Theme .....	76
4.3.5	Summary .....	88
4.4	Nu’s Emotional Experiences.....	88
4.4.1	General Background of Nu.....	88
4.4.2	The Research Setting.....	89
4.4.3	Nu’s Emotional Experiences with Lucy .....	89
4.4.4	The Emotional Theme .....	91
4.4.5	Summary .....	98
4.5	Lulu’s Cultural Experiences.....	98
4.5.1	General Background of Lulu .....	99
4.5.2	The Research Setting.....	99
4.5.3	Lulu’s Cultural Experiences with Lucy .....	99
4.5.4	The Cultural Theme.....	101
4.5.5	Summary .....	109
4.6	Nim’s Linguistic Experiences .....	109
4.6.1	General Background of Nim .....	109
4.6.2	The Research Setting.....	110
4.6.3	Nim’s Linguistic Experiences with Lucy .....	110
4.6.4	The Linguistic Theme .....	113
4.6.5	Summary .....	122
4.7	Num’s Relational Experiences .....	122
4.7.1	General Background of Num.....	122
4.7.2	The Research Setting.....	123
4.7.3	Num’s Relational Experiences with Lucy .....	123
4.7.4	The Relational Theme .....	125
4.7.5	Summary .....	131
4.8	Conclusion .....	131
<b>Chapter Five: Conclusions, Recommendations And Summary .....</b>		<b>133</b>
5.1	Conclusions for Research Questions.....	134
5.1.1	How did users feel when using Lucy to learn English?.....	134
5.1.2	How did users’ emotional responses to Lucy’s learning interface influence their cognitive processes? .....	136
5.1.3	At what level in the affective domain of learning did Lucy facilitate users’ learning? .....	137
5.2	Practical Implications.....	137
5.2.1	Designing Multimodal Communicative Language Learning Experiences .....	137
5.2.2	Designing a User – Adaptive Interface.....	139
5.2.3	Methodological Issues .....	139
5.3	Recommendations for Future Research .....	140
5.4	Summary .....	141
<b>Bibliography:.....</b>		<b>142</b>
<b>Appendix I: Participants Recruitment .....</b>		<b>151</b>

<b>Appendix II: Pre-Research Interview .....</b>	<b>152</b>
<b>Appendix III: Interview Questions .....</b>	<b>153</b>
<b>Appendix IV: Certificate of Approval.....</b>	<b>154</b>

## List of Tables

TABLE 2.1 : AXES OF EMOTIONAL SETS ( ADAPTED FROM KORT, KEILLY, & PICARD, 2001) .....	39
TABLE 2.2: THE MODEL OF EMOTION IN LEARNING (ADAPTED FROM KORT, KEILLY, & PICARD, 2001).....	39
TABLE 2.3: CIRCULAR AND HELICAL FLOW OF EMOTION (ADAPTED FROM KORT, KEILLY, & PICARD, 2001)	39
TABLE 2.4: AFFECTIVE DOMAIN: KNOWING (ADAPTED FROM KRATHWOHL, BLOOM, & MASIA, 1964) .....	46
TABLE 3.1: INSTALLATION OF CHATBOT LUCY .....	56
TABLE 3.2: CHATBOT LUCY'S CHARACTERS .....	58
TABLE 3.3: FIELD NOTES .....	63
TABLE 4.1: FIVE STAGES OF ANALYSIS.....	67
TABLE 4.2: USERS' EXPERIENCES - PFFECTIVE STATES .....	70
TABLE 4.3: USERS' AFFECTIVE STATES - AMOTIONAL RESPONSES .....	70
TABLE 4.4: USERS' EXPERIENCES - PMOTIONAL RESPONSES .....	71
TABLE 4.5: SENSE OF SIGHT AND HEARING - AFFECTIVE STATES .....	78
TABLE 4.6: AFFECTIVE STATES - AMOTIONAL RESPONSES .....	78
TABLE 4.7: SENSES - NMOTIONAL RESPONSES .....	78
TABLE 4.8: AXES OF LEE'S EMOTIONAL SETS IN LEARNING .....	79
TABLE 4.9: THE MODEL OF INTERPLAY BETWEEN LEE'S AFFECTIVE STATES AND HIS LEARNING CYCLE: ..	82
TABLE 4.10: SENSORY-AFFECTIVE DOMAIN:.....	85
TABLE 4.11: LUCY'S EMOTION GENERATION – NU'S AFFECTIVE STATES.....	94
TABLE 4.12: NU'S AFFECTIVE STATES – AU'S EMOTIONAL RESPONSES .....	94
TABLE 4.13: LUCY'S EMOTION GENERATION - NU'S EMOTIONAL RESPONSES.....	94
TABLE 4.14: AXES OF NU'S EMOTIONAL SETS IN LEARNING .....	95
TABLE 4.15: THE MODEL OF INTERPLAY BETWEEN NU'S AFFECTIVE STATES AND HIS LEARNING CYCLE: ..	96
TABLE 4.16: : EMOTIONAL-AFFECTIVE DOMAIN .....	97
TABLE 4.17: AXES OF LULU'S EMOTIONAL SETS IN LEARNING .....	104
TABLE 4.18: LULU'S CULTURAL BACKGROUND – CFFECTIVE STATES .....	105
TABLE 4.19: LULU'S AFFECTIVE STATES – AMOTIONAL RESPONSES .....	105
TABLE 4.20: LULU'S CULTURE – LMOTIONAL RESPONSES .....	105
TABLE 4.21: THE MODEL OF INTERPLAY BETWEEN LULU'S AFFECTIVE STATES AND HER LEARNING CYCLE:	107
TABLE 4.22: : CULTURAL-AFFECTIVE DOMAIN .....	108
TABLE 4.23: LINGUISTIC – NFFECTIVE STATES .....	115
TABLE 4.24: AFFECTIVE STATES – AMOTIONAL RESPONSES .....	115
TABLE 4.25: LINGUISTIC – NMOTIONAL PROCESSES.....	115
TABLE 4.26: AXES OF NIM'S EMOTIONAL SETS IN LEARNING .....	119
TABLE 4.27: THE MODEL OF INTERPLAY BETWEEN NIM'S AFFECTIVE STATES AND HER LEARNING CYCLE:	119
TABLE 4.28: LINGUISTIC-AFFECTIVE DOMAIN .....	121
TABLE 4.29: RELATIONAL – LFFECTIVE STATES.....	127
TABLE 4.30: AFFECTIVE STATES – AMOTIONAL RESPONSES .....	127
TABLE 4.31: RELATIONAL – LMOTIONAL PROCESSES.....	127
TABLE 4.32: AXES OF NUM'S EMOTIONAL SETS IN LEARNING .....	128
TABLE 4.33: THE MODEL OF NUM'S AFFECTIVE STATES IN LEARNING.....	129
TABLE 4.34: RELATIONAL-AFFECTIVE DOMAIN .....	130
TABLE 5.1: USERS' AFFECTIVE STATES .....	135
TABLE 5.2: USERS' EMOTIONAL RESPONSES .....	135

## Lists of Figures

FIGURE 1: LUCY'S FIVE LEARNING MODULES .....	57
FIGURE 2: SAMPLES OF LUCY'S TOPICS .....	57
FIGURE 3: LUCY'S TRANSLATED LANGUAGES .....	58
FIGURE 4: FOUR COMPONENTS OF EACH LUCY'S MODULE .....	114

## Preface

Over 4 years ago, as a new instructor at Beijing University of Chemical Technology, I taught an introductory class on college English Listening and Speaking. As a non-native speaker, I felt challenged to provide feedback to students' pronunciation problems; as a language teacher for an introductory language class, I felt like an emotional laborer (Hochschild, 1983, p. 7). I was expected by my students to behave in a patient way and to be pleasant as well as obliging. I had to hide my true feelings towards thousands of pronunciation repetitions and express my shaped feelings in front of my students. I had to manage my emotions to “create a publicly observable facial and bodily display” (Hochschild, 1983, p. 7). I became an emotional laborer.

For one meeting, early in the semester, I assigned the CSIEC<sup>1</sup> (Computer Simulator in Educational Communication), a chatbot system, which was installed on an http server in China to help Chinese students learn English. It was advertised in ten famous Bulletin Board Systems (BBS) of Chinese universities claiming that users could chat with the system in English: a learning partner of foreign languages.

I began class by reflecting on the learning outcomes of chatting with the CSIEC. To my surprise, my students, who were mostly younger than 30 years old, instead of talking about the learning outcomes, wanted to share how they treated the chatbot as a friend, as a puppet or as a pet and how they felt a connection with the chatbot such as excitement, interest, motivation, frustration and boredom. The learning outcomes were not optimal due to the technical issues of the chatbot technology at times. For example, the CSIEC

---

<sup>1</sup> <http://www.csiec.com/>



mostly repeated its responses, responded with irrelevant topics of the dialog and showed apathy to students' frustration and boredom, which discouraged them from continuing the conversation. However, students found that their emotional attachment to the CSIEC shaped their willingness to use this technology for their language learning. As a result, although I stopped using this innovative technology in teaching my introductory language class, I did not stop my passion in researching multimodal language education interface design, which could combine a visual modality (a display, keyboard, and mouse) with a voice modality (speech recognition for input, speech synthesis/recorded audio for output). I believed that the future multimodal adaptive language education interface design could benefit from this investigation, and this innovative multimodal adaptive interface could promote language education in powerful ways as well as help teachers who performed emotional labor to feel emancipated from institutionalized emotions (Hochschild, 1983).

Inspired by the innovative research project: "Emotech: Emotion and Technology" initiated by my supervisor, Dr. Stephen Petrina, at the University of British Columbia, I decided to devote my thesis to the study of multimodal adaptive language education interface designs that employ chatbot technology. Rather than focusing on technical design details, I shed light on users' (learners') everyday experiences with the existing language education chatbot technology, "Lucy", hoping to investigate decisive factors that influence multimodal adaptive language education interface design.

## **Acknowledgements**

The completion of this thesis is only possible due to the support and encouragement of many people. First of all, I would like to thank my supervisor, Dr. Stephen Petrina, for his tremendous help with my research design, data collection, and thesis writing, especially his emotional support. Without his help, my thesis is impossible. I would like to thank Dr. Don Krug for his incisive comments on my research design. I thank Dr. Sidney Fels with his support for my Human – Computer Interaction application and completion as well as his help with this study.

I appreciate helps from my colleagues Lauren Hall, Dai Kojima and Juyun Kim in the New Media Lab and Bonnie Wen for their support of the data collection and thesis defense.

I also would like to give my thanks to Basia Zurek and Bob Hapke, who have been providing me lots of help for my completion of the program.

I thank my husband Jigang Wang for his continuous support of my studies at the University of British Columbia. He is always reliable, responsive and helpful with my career and education. I also thank my parents and Jigang's parents for always supporting me throughout my education.

## Chapter One: Introduction

Teacher: Please tell me about learning outcomes of communicating with the CSIEC.

Student A: It was exciting at the beginning. I felt he was a real person. But when I continued to talk to him, he couldn't follow my conversation.

Student B: I don't think I can learn English from her. However, I would say she might be my friend who could talk to me in English. She can't follow my talk, but it is fine. At least, she can listen to me and tolerate my weird pronunciation.

Student C: To me, I think he was a learning companion... Although he can't correct my grammar mistakes, pronunciation and intonation problems as well as spelling errors, I feel a connection with him.

Student D: I couldn't learn. I got a problem of logging into the website. I don't know how to fix it.

Teacher: How long did you spend talking to the CSIEC?

Student A: Twice a day. I started the communication in the early morning before I went to class; I started again in the afternoon before I went to the dining hall. Every time it lasted 30 minutes.

Student B: I don't have a regular time to talking to the CSIEC, sometimes twice a day, sometimes once a day and sometimes none. It depends on when I am available.

Student C: I tried it once a day in the morning before I went to class.

Student D: I only tried it once after you assigned the CSIEC to us. And I found I didn't know much about computer and couldn't log in, so I quit.

Teacher: How did you feel when communicating with the CSIEC?

Student A: It is hard to say. I chose a very cute character and it was impressive.

Student B: I feel comfortable not talking to a language teacher during my language learning processes. I treated her as a friend who lived inside the machine and spoke English.... It was a wonderful experience to have her.

Student C: For me, I know that he won't point out my errors, so I feel more relaxed when talking to him. Just like talking to myself in the mirror in order to practice English, I got another myself in the machine.

Student D: I don't know. I spent lots of time trying to log in and failed. I don't know the reason and I can't find any help, so I quit the program.

Teacher: How did you like the CSIEC?

Student A: I like the idea of using computers to help learners practice English, but I don't think the CSIEC is well-designed language tutor. He can't replace a human teacher.

Student B: I like her, but I don't think she can become my language teacher. She doesn't know how to teach, but it doesn't matter, for learning can happen among peers. One of advantages of the CSIEC is that it is free.

Student C: He is always friendly. That is amazing... My mood often changes when I am practicing the pronunciation. I appreciate that he listens to me all the time.

Student D: I don't like it. I would rather go back to the classroom and learn from teachers. I hate technology. It makes me confused and wastes my time on troubleshooting technical problems.

This was from conversations I had with my students four years ago when I first used technology in my English Listening and Speaking Class. Although the learning

outcomes were not what I expected – students couldn't learn much from the CSIEC system. The CSIEC system took them away from what they found meaningful for language learning, but they still acknowledged that the CSIEC learning system did enrich their learning experiences. When using the CSIEC, their affective states and emotions frequently shaped and colored their learning experiences and became central parts of those experiences.

Therefore, I regarded the CSIEC, a chatbot system, which is a computer program designed to stimulate an intelligent conversation with one or more human users, as an amazing potential technology to cope with users' increased demands of language practice and possibly free language teachers from being emotional laborers. I discussed chatbot technology in this study and conducted an ethnographic inquiry into users' experience with chatbot technology. A sensory theme, emotional theme, cultural theme, relational theme and linguistic theme are five key components to designing a multimodal adaptive language education interface.

## **1.1 *Research Rationale***

“We don't just use technology; we live with it ” (MaCarthy & Wright, 2004, p. ix). The depersonalized drill and practice interface leaves the user bored and unable to complete a simple learning task. Is it possible to design language learning interfaces that are more usable? Long gone are the days when language learning interfaces are seen as merely powerful cognitive tools or media, which are devoid of users' motivation to use and hinder them to reach their desired learning goals. Much more deeply than ever before we are aware that end-users influence how a design takes shape and users have a deep impact on the design. Interacting with technology involves users emotionally,

intellectually and sensually” (MaCarthy & Wright, 2004). Users experience technology in accordance with their emotions and intellect. They become a central part of the development process and are placed at the center of the design. As a result, the role of the designer is to facilitate the task for the user and to make sure that the user is able to make use of the product as intended and with a minimum effort to learn how to use it (Abrams, Maloney, & Preece, 2004). Hence, people who design interactive systems need to be able to understand and analyze users’ experience with technology (Norman, 2006). Therefore, this research contributed to the study of the need to fully explore users’ experiences with language education interface during design processes.

Secondly, the increasing computational power helps users complete tasks more efficiently by shifting the burden of adaptation on users to that of the computer.

We are entering an era of “new computing”: The old computing was about what computers could do; the new computing is about what users can do. Successful technologies are those that are in harmony with users’ needs. They must support relationships and activities that enrich users’ experiences. (Shneiderman, 2002, p.2)

Thus, in this study, I paid attention to the adaptive user interface, which can provide a flexible mechanism for systems to adapt to the needs of different users for a variety of tasks. Such systems are aware of and able to adapt to the particular needs of specific users.

Thirdly, It is no longer sufficient for a product to be simply usable or aesthetically pleasing, but it needs to evoke positive emotional responses (Pace, 2004; Romaine, 1994). Users’ feelings can’t be separated from their cognitive capacity when using technology.

Technology can't be simply viewed as a cognitive tool – a tool that stored and manipulated data in ways far beyond human capacity, but should be laden with emotional capacity. A positive emotional learning environment makes users feel confident in using the learning interface and derive pleasure or joy from the interface. Otherwise, fear would be generated in some users when they were confronted with learning systems or considerable distress would occur, which might get users trapped in a “loop of fear-failure-decreased motivation” (Petrina, 2007, p. 67). Hence, A reliable language learning interface design for “reducing negative emotions such as fear, envy, and anger and increasing empathy, sympathy and pleasure” should be developed (Petrina, 2007, p. 67).

Last but not the least, the past two decades have seen a powerful growth of speech-driven user interfaces. The widespread acceptance of speech as a human-computer interface supports new training systems for learning foreign language such as chatbot CSIEC and chatbot Lucy. As applications generally have become more complex, a single modality does not permit the user to interact effectively across all tasks and environments (Larson, Oviatt, & Ferro, 1999). A multimodal interface offers the user freedom to use a combination of modalities or switch to a better-suited modality, depending on the specifics of the task or environment (Oviatt et al., 2000).

## **1.2 Research Purpose**

This research approached user-centered speech-driven multimodal adaptive language learning interface design. In order to achieve this goal, I shed light on investigating user's experiences with the commercial speech-driven chatbot Lucy's interface.

Initially my idea was to use the CSIEC to explore user-centered speech-driven multimodal adaptive language learning interface design, but soon I noticed the replace of the commercial chatbot Lucy, a digital language tutor that could carry on extensive conversation with users as they speak to their computers through a microphone. Using the advanced speech recognition system, Lucy could communicate with its users in English and guide them through useful exercises to improve their pronunciation. So I employed Lucy and conducted an ethnographic inquiry of users' experiences with Lucy hoping to investigate decisive factors that influence multimodal adaptive language learning interface design.

I empirically analyzed five factors: learners' sensory theme, emotional theme, cultural theme, relational theme and linguistic theme. These five themes formed learners' experiences with Lucy's learning interface and mediated their cognitive processes when using the chatbot. I argue that these five themes reflect how the learners interact with the interface. "In this respect, multimodal interfaces have the potential to accommodate a broader range of users than traditional graphical user interfaces (GUIs) and unimodal interfaces" (Oviatt et al., 2000, p. 270). Finally, a speech-driven multimodal adaptive language learning interface design recommendations and future research directions are provided.

The scientific virtues of reduction and generalization were dominant approaches to understand relationships between people and technology since the 1970s and the early 1980s. "In this context, the computer was seen as a tool through which set work was accomplished" (MaCarthy & Wright, 2004, p. 6). In the late 1980s and throughout the 1990s, the user was recognized as a social actor (MaCarthy & Wright, 2004) and the



focus was shifted to the contingent aspect of everyday activity and “was geared toward asserting the salience of the social context of activity in discourse about people and technology” (MaCarthy & Wright, 2004, p. 7). This movement insisted that “all action was richly contextualized” (MaCarthy & Wright, 2004, p. 8). Finally the 1990s witnessed the computer becoming a consumer product. Research and industry were paying much more attention to the users’ experiences. As a consequence, “interaction with technology is now as much about what people feel as it is about what people do” (MaCarthy & Wright, 2004, p. 9). As a result, users’ research methodology turned to ethnographic inquiry.

Hence, in this study, I conducted a laboratory ethnographic inquiry on users’ experiences with the chatbot Lucy in two labs: one lab at an English Language Institute at a University and the other lab in the education building at a University. This ethnographic inquiry included 1) observing the users’ experiences with the chatbot Lucy; 2) analyzing users’ emotional responses to the chatbot Lucy’s interface in the service of their cognitive processes; 3) analyzing affective learning levels that the chatbot Lucy could achieve; and 4) providing language education interface design recommendations and future research directions.

### ***1.3 Research Problems***

We live with technology, however, “there is little history of interest in user experience in HCI and related research areas” (MaCarthy & Wright, 2004, p. 5). Most HCI approaches neglect or at least underplay the role of the personal experiences users have in interacting with technology.

HCI and related disciplines are not used to dealing with (users') experiences. HCI grew out of collaboration between the disciplines of computer science and psychology and was directed more toward functional accounts of computers and human activity than toward (users') experiences. Against this background, it might be worth looking briefly at the emergence of interest in (users') experiences with technology and how HCI currently understands users' experiences. (MaCarthy & Wright, 2004, p. 6)

The users' role in HCI is changing: cog in a rational machine in 1970s, a source of error in the 1980s and then a social actor in the 1990s and now a consumer (Kuutti, 2001). In order to design a user-centered interface, we cannot neglect the wide range of influences of user's experience with technology (MaCarthy & Wright, 2004).

There has been a debate in the computing community between those who promote intelligence in the interface and those who promote comprehensible, predictable, and controllable interfaces that give users a sense of power, mastery, control and accomplishment (Kaufman & Weed, 1998). The former refers to what is called adaptive interface, which relies on intelligence to automatically adjust in a way that is expected to better suit the needs of each individual user and the latter refers to what is called adaptable interface, which, in contrast, keeps the user in control by providing the mechanisms for the user to personalize according to his/her needs (McGrenere, Baecker, & Booth, 2002). The debate influenced two research directions on emotions in the computing world: people express their emotional states through/with computers (adaptable), which is known as research in user experience; and computers address/respond/recognize users' emotions and adjust their interactions with users

(adaptive), which is known as affective computing. This research investigated users' experiences – how users express their emotional states through/with computers in order to examine the benefits of an adaptive learning interface.

Thirdly, the current human–computer interaction interfaces are almost exclusively unimodal interfaces, “where there is only a single channel for data output (e.g. screen) and only a single channel for data entry (e.g. keypad)” (Ringland & Scahill, 2003, p. 182). However, if multiple channels (or modes) are used for input and output such as spoken data entry using speech recognition technology or screen output with audio, a much richer user interface can be created. There are many options that are possible nowadays, like eye tracking or gesture recognition (sign language). This research focused on two modes, specifically a voice interface combined with the traditional graphical user interface.

## **1.4 Research Questions**

In order to investigate user–centered speech–driven multimodal adaptive language learning interface design, three specific questions were formulated and addressed:

Question 1:

How do users feel when using chatbot Lucy to learn English?

To understand users' experiences with chatbot Lucy, I investigated users' perceived main weaknesses of existing language educational interfaces such as Lucy; how the multimodal learning interface, which combined voice interface and graphic user interface, influenced users' feelings; How Lucy's inability to address users' emotions made them feel.

By employing the think aloud method – verbal reports on mental processes, I adopted participant reconstructive interviews to allow my participants to articulate their experience with the chatbot interface. However, users' emotional responses to Lucy was not enough to research user-centered multimodal adaptive learning interface design. This learning interface needed to facilitate learning in the first place; hence, users' experiences in its service of cognition was also investigated. Thus, the second question posed for this study was:

Question 2:

How do users' emotional responses to the chatbot Lucy's learning interface influence selected cognitive processes?

The second question aimed to examine whether Lucy facilitated learning or actually hindered the progress of learning. Understanding users' learning processes in this study involves investigating cognitive processes.

Additionally, the learning interface included a planned instructional objective embedded in its learning model. The objective determined in advance what the learning model wanted users to know, feel or do, or how it wanted them to act or behave. Lessons or contents could be developed to lead to intended results (Petrina, 2007). The traditional practice and drill language learning unimodal interface was typically focused on the cognitive domain and psychomotor domain of learning, but was devoid of the affective domain of learning. However, a successful language learning interface ought to cover all three instructional domains of learning – cognitive domain, affective domain and psychomotor domain (Petrina, 2007). Therefore, research into users' learning with Lucy ought to include the affective domain with the purpose of exploring if Lucy could help

users (learners) achieve higher levels of the affective domain (valuing, organization & characterization). Therefore, a third question was addressed:

Question 3:

At what level in the affective domain of learning does Lucy facilitate users' learning?

Through examining learning objectives embedded in the affective domain in Lucy's learning models, this third question investigated how higher level of objectives can be designed into language educational interfaces in order to provide better support for student performance.

### **1.5 *Limitations of The Study***

Due to the limited amount of time and other practical reasons, there are several limitations associated with this study. The first limitation of this study related to its design. Since the data collection was limited to a university, and only five users' experiences were analyzed, the findings of this study might lack transferability to other users in similar educational settings. In addition, the interaction recorded focused on a lab research design, which differed from the real practice either at home or in classroom. The participants might have been uncomfortable with videotaping, but they nonetheless displayed a range of feelings related to the interaction with Lucy.

Some of the users in this study showed a great interest in interacting with Lucy, however, after two or three interactions, they quit from the study. Fifteen users participated with study and one limitation was the selection of five cases for analysis (see Section 4.2).

A possible limitation in terms of data collection related to the techniques employed, which were described in Chapter Two. For example, some parts of the setting were

incomplete due to the camera angle. Another limitation came from the overlapping users' experiences during the interaction. In order to emphasize one of experiences that users experienced, I had to distinguish among users' different experiences in the process of analysis. It was especially difficult when the emotional word stated by users had multiple referrals. For example, when Nu described his feelings toward speech recognition system, the emotional word not only showed his affective experiences but also involved his hearing engagement (sensory). A further limitation of this study was that the interview data were not directly linked to users' experiences during the interaction. Although video as a stimulus was employed during the interview, users constructed their feelings and experiences based on recall, which sometimes differed from what I observed and what was from the think aloud. But the interview did assist understanding the perceptions of the participants about their performance during the interaction.

## **1.6 Thesis Structure**

To achieve my goals stated above, I divided the thesis structure into five chapters. Chapter one is an introduction of the research background, purposes, problems and questions as well as limitations. Chapter two is a critical review of the literature on chatbot technology, multimodal interface, adaptive interface, Evans' definition of emotion, Kort's emotion and learning theory, as well as Krathwohl's affective domain of learning. This chapter builds on the theoretical framework of Evans' definition of emotion, Kort's emotion and learning theory as well as Krathwohl's affective domain. Chapter three acquaints readers with an elaborated account of the research design for the study of user-centered speech-driven multimodal adaptive language learning interface design. Chapter four provided multi-layered analyses of five users' stories about their

experiences with Lucy's interface. Five pivotal themes: the sensory theme, emotional theme, cultural theme, linguistic theme and relational theme are discussed. Based on analyses in Chapter four, Chapter five is an attempt to provide recommendations on user-centered speech-driven multimodal adaptive language learning interface design. The findings and implications are discussed and future directions for researching multimodal adaptive language learning interface design are proposed.

## **Chapter Two: Literature Review**

The past decade has witnessed a great deal of interest in technology driven language learning. Various technologies have been employed in many settings and environments to provide assistance to language learners. Because of advances in pedagogy and learning paradigms, the traditional behaviorist language learning model has shifted to a communicative and integrative model, and thus, in accordance with this shift, the traditional unimodal language learning interface has moved towards a multimodal communicative and adaptive interface in CALL (computer assisted language learning) systems.

In this chapter, I argue that user-centered speech-driven multimodal adaptive language learning interface would further enhance CALL systems and the integration of multimodal and adaptive interface in CALL systems would benefit from a new generation of CALL applications. In the following sessions, I present a brief history of CALL's three main stages: behaviorist, communicative and integrative CALL and then discuss chatbot technology for enhancing language learning situations. Adaptive interface, especially affective computing, is discussed; debates in the field of adaptive interfaces are raised. Following the adaptive interface, the benefits of multimodal technologies that could be used in CALL systems are described. This chapter concludes with a discussion of Kort's learning model and Scheffler's cognitive emotion theory as well as Krathwohl's affective domain of learning.



## **2.1 *The History of Computer – Assisted Language Learning***

Three main stages in the history of CALL include behaviorist, communicative, and integrative CALL (Warschauer & Healey, 1998). The behaviorist stage, resting on the behaviorist learning model, is a sub-field of computer-assisted instruction. These interfaces are considered as tireless tutors that provide users with endless drills and exercises. However, the heavy criticism of the behaviorist approach, pedagogically as well as theoretically, prompted the communicative CALL approach to emerge in the early 1980s. The communicative CALL approach argues that learning is an explorative process, and the *use* of language should be the focus. But the criticism of this communicative approach from socio-cognitivists during the 1990s demands an even higher degree of language use: the use of language should take place in authentic social contexts. This shift puts forward an integrative model where connected multimedia and novel technologies such as chatbot systems are readily available. Although the categorization of CALL applications according to these three stages is simplified, it serves the purpose of understanding contemporary development of CALL. Based on the overview of Warschauer and Healey (1998), I believe that multimodal practice with feedback and individualization in a large class are two of the main benefits of adding computer components to language instruction, which can be of great value to future CALL applications: multimodality and adaptivity.

## **2.2 *Chatbot Technology Overview***

### **2.2.1 Chatbot technology development**

Chatbots are known by a wide variety of terms including chatterbots, virtual assistants, virtual agents, intelligent agents, or web-bots. I prefer chatbot as it is the

simplest and most commonly understood term. Wikipedia defines chatbot as a computer program designed to simulate an intelligent conversation with one or more human users via auditory or textual methods. Chatbots are simple to use; users chat through text or voice input over a computer screen with chatbot's output or audio/voice output. With embedded speech recognition systems, the chatbot recognizes users' voice inputs, consults its knowledge base or programming language, synthesizes voice and replies. The conversation continues as long as it is interesting or useful for the user.

Previous work in the chatbot area dates back to the first well-known chatbot, ELIZA, which was designed by Joseph Weizenbaum and released in 1966 (Weizenbaum, 1966). ELIZA parodied a Rogerian therapist, largely by rephrasing many of the patient's statements as questions and posing them to the patient, such as typing in the word "mother" would cause ELIZA to respond "tell me more about your family." The ELIZA chatbot became a point of reference for other programs using similar techniques for providing a conversational interface.

After ELIZA came PARRY, written by psychiatrist Kenneth Colby in 1972 at Stanford University. PARRY attempted to simulate a paranoid schizophrenic and was modeled on the paranoid mind. Many expert psychiatrists found it difficult to tell whether PARRY was human or not. The program was designed to emit linguistic responses based on internal affective states. To create this effect, three measures – fear, anger and mistrust – were used and their values changed depending on the flow of the conversation.

Later at Carnegie Mellon University, Dr. Richard Wallace developed A.L.I.C.E., whose brain was inscribed in an XML-based language called AIML (Artificial Intelligence Markup Language). A.L.I.C.E. was the winner of the 2000 Loebner

competition and “her brain” engaged in a conversation with a human user by applying categories containing a stimulus, or pattern, and a template for the response. “Category patterns were matched to find the most appropriate response to a user input. Further AIML tags provided for consideration of context, conditional branching a supervised learning to produce new responses” (Kerly, Hall, & Bull, 2007, p. 177).

British programmer Rollo Carpenter created the Jabberwacky chatbot, which learned continuously from its conversations with web users. The Jabberwacky chatbot winner, George, embedding chatbot speech and voice recognition, enabled chatbots whose avatars went beyond the Oddcast SitePal technology, common in online Pandorabots, by displaying emotional body language to suit the topic and general mood. The technology behind Jabberwacky worked on a different principle to that of other artificial intelligence software being developed. The system learned from all its previous conversations with human users. There were no fixed rules or principles programmed into the system and it operated entirely through user interaction. The system stored everything that was said to it and used contextual pattern matching techniques to select the most appropriate response. “It had no hard-coded rules, instead relying entirely on previous conversations. It was explicitly not intended to do anything useful, instead being simply to chat” (Kerly et al., 2007, p. 179). Therefore, the program created a massive database of contextually appropriate conversations and chose an appropriate response it learned from a previous user when holding a conversation.

“Modern commercial chatbots, such as those developed with Lingubot technology, offered sophisticated development environments allowing the building of intelligent conversational agents with complex, goal driven behavior” (Kerly et al., 2007, p. 178).

The underlying technology was based on a sophisticated word and phrase pattern recognition system that matched pre-programmed responses in the Lingubot's knowledge base with questions typed in by users. "In Lingubots' both the words and the grammatical structure of the user's input were analyzed using customized templates. This facilitated the development of a user model, which was used in conjunction with the conversational context and specific words, in the dialogue to determine the chatbot's response. Responses might include further conversation with the user, reading or writing to external systems or a combination of these. This rich range of responses allowed for intelligent conversation with the user, and provided the ability to steer the user back to the task in hand if they strayed from the designated discussion content for too long" (Kerly et al., 2007, p. 178).

### **2.2.2 Chatbot technology usage**

The idea of a chatbot has been around for a long time. They are now generally emerging from the universities and research laboratories and becoming ready for common use. There are hundreds of different chatbots, developed for a variety of reasons; they range from hardwired programs with simply coded patterns to systems built upon embedded learning algorithms, which continuously expand their language knowledge base. Chatbots are created for fun such as virtual characters and entertainers, or as part of interactive games such as game player, or designed to provide specific information and direct dialogue to specific topics such as website guide, frequently asked questions (FAQ) guide, virtual support agent, virtual sales agent, survey taker, quiz host, learning tutor and chat-room host.

Among hundreds of different uses of chatbot, the tutor role played by chatbot is proving to have a significant impact in the CALL area. Many traditional CALL technologies have fallen short of expectations because of the relatively sterile nature of the experience – there is no bond between student and computer in the way that there is between student and teacher. However, chatbot technology as a learning tutor provides a way to recreate that student – teacher bond, giving the user a character to empathize with – increasing their enjoyment and commitment to language learning. The chatbot can take on a simple mentoring role – offering encouragement and general learning advice, or it can be provided with the course knowledge and answer questions that a student may have.

### **2.3 *Adaptivity***

Adaptivity refers to a technological system's ability to dynamically adjust its behavior and settings to an individual user at use-time. This adaptive ability of technology represents a crucial role in catering for individual differences in all learning situations and can help solve a problem in large language classes where teachers are trying to accommodate all needs of different learning pace students.

However, there has been a debate in HCI community between those who promote intelligence in the interface and those who promote comprehensible, predictable, and controllable interfaces that give users a sense of power, mastery, control and accomplishment (Kaufman & Weed, 1998). The former refers to what is called adaptive interface, which relies on intelligence to automatically adjust in a way that is expected to better suit the needs of each individual user and the latter refers to what is called adaptable interface, which, in contrast, keeps the user in control by providing the mechanisms for the user to personalize according to his/her needs (McGrenere et al.,

2002). Hence, some researchers who represent an affect-proponent viewpoint argue that the ability of emerging technologies to address users' affect is a key component of what constitutes effective and desirable HCI (Minsky, 2006; Picard, 1997). And others who represent an affect-skeptic viewpoint argue that "affective computing is neither a meaningful concept nor a reasonable goal" (Hollnagel, 2003, p. 70). They refute the proposition that technology should have the ability of recognizing emotions by arguing that "rather than trying to make computers (or computing) affective, we should try to make communication effectual. Rather than trying to reproduce emotions, we should try to imitate those aspects of emotions that are known to enhance the effectiveness of communication" (Hollnagel, 2003, p. 70). As a result, this debate generates two research directions on emotions in the computing world: people express their emotional states through/with computers (adaptable), which is known as research in users' adaptable experiences; and computers address/respond/recognize users' emotions and adjust their interactions with users (adaptive), which is known as affective computing.

Spanning all digital technologies in the learning environment, long gone are the days when educational technology can be seen as only a powerful cognitive tool or medium, proving to the learner to be a non-judgmental and patient tutor. "The burden of adaptation has gradually been shifting from the human user to the computer" (Hudlicka, 2003, p. 2). As the range of computer applications broadens, the topic of the decreasing tolerance of user frustration drives the human-computer interface more towards adaptivity (Hudlicka, 2003; Lyytinen & Yoo, 2002).

Secondly, since more and more CALL applications aim at moving away from pure behaviorist drills towards a combination of behaviorist drills and the understanding of the

complexity of the use of language, an ability of technology to recognize learners' emotions becomes important. Researchers examined human language tutoring and found that expert human tutors devoted at least as much time and attention to the achievement of affective and emotional goals in tutoring as the achievement of the sorts of cognitive and informational goals (Lepper & Chabay, 1988). It is difficult to deny the role of affect in language learning such as motivation, interest or frustration. Hence, there is a need to develop interfaces, together with new signal processing, pattern recognition, and reasoning algorithms for assessing and responding to the affect of the learner in real time.

Thirdly, the mounting evidence of the importance of emotions in human-human interaction provides the basis for researchers in the engineering, HCI and computer science communities to develop automatic ways for computers to recognize emotional expressions as a goal towards achieving human-computer intelligent interaction. The question is not whether intelligent machines can have emotions, but whether machines can be intelligent without any emotions (Minsky, 1988).

One of the problems with studying emotions in technology is what it is. No topic of our mental life seems more complicated, fascinating and important to us than emotions. Since William James (1884) asked the question "What is an emotion", the ontology of emotion continues to baffle us. Researchers and theorists in different disciplines have battled for the soul of emotions for thousands of years. Little agreement and research trends toward a definition of emotion shape affective HCI towards initiating many state-of-the-art research projects based on different emotion traditions and inventing innovative technologies to address emotions.

In the following section, I first review definitions of emotions, which are categorized by different research traditions and then discuss the definition of emotion proposed by Evans. Finally, I describe current technologies used in emotion recognition and research projects done in this field.

### **2.3.1 The ontology of emotion – Emotion research traditions**

Throughout history, the terms “affect,” “emotion,” “feeling” and “mood” are often used interchangeably (Batson, Shaw, & Oleson, 1992; Cohen & Areni, 1991; Derbaix & Pham, 1998). Nearly a hundred definitions of emotion have been categorized since the 1980s with some favored by different disciplines and research traditions. Five main theoretical traditions, Darwinian-evolutionary theory, body reaction, cognitivism, behaviorism and social constructivism, provided their own explanation of understanding the ontology of emotion.

The body reaction theory has been prevalent since Ancient Greece and has largely been taken for granted in emotion theory until the beginning of the 20<sup>th</sup> century. Emotion within this paradigm is seen as feelings caused by changes in physiological conditions relating to autonomic and motor functions (James, 1884). Emotion is “a special state of consciousness or bodily state” (Scarantino, 2005, p. 8). For example, when we perceive that we are hurt, this perception sets off a collection of bodily responses such as increased heart beat, blood pressure, rate of respiration, and gastric activity, decrease in saliva flow, trembling and etcetera and our awareness of these responses is what constitutes fear or anger. As what James put it: “we feel sorry because we cry, angry because we strike, afraid because we tremble, and it is not that we cry, strike, or tremble, because we are sorry, angry, or fearful, as the case may be” (James, 1884, p. 190). Thus, this tradition



treats different emotions as products of different patterns of autonomic response (James, 1950; Schachter, 1957). Aristotle, Descartes, Hume (1739), Freud, James (1884), Damasio (1994) and Prinz (2004) exemplified this theory.

However, opponents such as Walter Cannon (1929), Stanley Schacter and Jerome Singer (1962) argued that this body reaction theory was unable to give an adequate account of the differences between emotions. Schacter and Singer (1962) concluded that the differentiation of specific emotions was not physiological, but cognitive or something else. Another problem with body reaction theory is that it fails to account for emotions' various ties to rationality. This problem was to some extent mitigated when Damasio elaborated a sophisticated view of emotions that involved a capacity for the brain to monitor the body's past and hypothetical responses, but it failed to fully explicate the intentional nature of emotions.

In the late nineteenth century, Darwinian-evolutionary theory emerged and regarded emotion as adaptations whose purpose is to solve basic ecological problems facing organisms (Frank, 1988; R. Plutchik, 1979). In another words, it is "an adaptive solution to a fundamental life task" (Scarantino, 2005, p. 71). Pioneered by Darwin (1872), such an evolutionary framework to understand emotion was concerned not so much with the question of how our emotions might have evolved, but rather with why they should have the forms of expression that they have (Darwin, 1872). According to Darwin, emotional expressions once served particular functions but now accompany particular emotions because of their usefulness in communicating these emotions to others (Darwin, 1872). This evolutionary framework of emotions later was received by Silvan Tomkins (1995)

and Robert Plutchick (1962) and was brought to fruition in the 1960s by Paul Ekman (1969) and Carrol Izard (1971).

Inspired by Darwin, since the early 1970s, Paul Ekman and his colleagues have performed extensive studies of human facial expressions. They found evidence to support universality in facial expressions. Through studying facial expressions in different cultures, including preliterate cultures, they discussed the universal facial expressions including **happiness, sadness, anger, fear, surprise and disgust**. Their research found much commonality in the expression and recognition of emotions on the face. But they also observed differences in facial expressions. Their research results – when watching the same film clips, Japanese subjects and American subjects tended to show similar facial expressions with researchers' not presence in the room, but Japanese subjects were more reluctant to show their real expression with researchers' presence in the room, which showed that the culturally determined 'display rules' always arrived on the scene after the basic emotional response had been set in motion. Therefore, they proposed that facial expressions are governed by "display rules" in different social contexts (Ekman, 1972).

However, the view that emotion was essentially "a special way of dealing with fundamental life tasks" (Scarantino, 2005, p. 8) and different social and psychological emotional functions were shaped relatively independently by natural selection (Cosmides & Tooby, 2000) ignored those emotions that involved higher cognitive processes, such as jealousy and envy. This problem led many emotion researchers to stress cognitive aspects of emotions.

Cognitivist theory, characterizing emotions primarily in terms of associated cognition, was anticipated by the Stoics and followed by Robert Solomon (1980), Jerome Neu (1977) and Martha Nussbaum (2001). Under this tradition, emotion is described as judgments, sets of beliefs and desires or affect-laden judgments (Broad, 1954) and as complexes of beliefs, desires, and feelings (Oakley, 1992). The psychological appraisal theories under this tradition gained contemporary attention and were developed during the 1960s and 1970s by philosopher Anthony Kenny (1963) and Errol Bedford (1957) and psychologists Magda Arnold (1960), Stanley Schachter and Jerome Singer (1962). Magda Arnold (1960) first introduced the notion of appraisal into psychology, characterizing it as the process through which the significance of a situation for an individual is determined. Appraisal theories attempted to define emotion as “a special way to appraise” (Scarantino, 2005, p. 8), and argued that emotion was the integration of thought and perception with arousal (Russell, 1960; Schachter, 1964; Schachter & Singer, 1979; Sully, 1902). Appraisal theories generally allowed for cognitive processes underlying emotion to be either conscious or unconscious, and could involve either propositional or non-propositional content, but cognitivists typically claimed that emotions involved propositional attitudes. Robert Solomon (2003) and Martha Nussbaum (2001) in philosophy and Richard Lazarus (2001) and Klaus Scherer (2001) in psychology continued to offer updated versions of this approach.

However, this theory faced criticism along a number of fronts: first of all, it was criticized because the view of emotions as propositional attitudes excluded animals and infants lacking language (Deign, 1994). The second criticism was known as “a fear of

flying” objection: I may be well aware that flying is the safest means of transport and yet suffer fear of flying (Stocker & Hegeman, 1992).

The ontological question of emotion is also addressed by behaviorism, which regards emotion as “a special disposition to behave” (Scarantino, 2005, p. 8). That is, if anger was defined, it would mean “a disposition to attack the object of anger” (Scarantino, 2005, p. 8). This theory was developed at the beginning of the twentieth century by psychologist John Watson and then further received by B. F. Skinner in psychology and Gilbert Ryle in philosophy. Although this theory collapsed around the mid-1950s, traces of a behaviorist understanding of emotion still existed in several contemporary theories, especially Frijda’s (1986) account of emotion as action tendencies<sup>2</sup>.

Lastly, social constructivism argues that emotion is “a special way of playing a social role” (Scarantino, 2005, p. 8). Thus, anger is “a social role in which one engages when wanting to be justified in the exercise of aggression” (Scarantino, 2005, p. 8). Anthropologists such as Catherine Lutz, philosophers such as Rom Harre and psychologists such as James Averill believe that emotions are essentially culturally specific social roles. Psychologist Brian Parkinson and philosopher Paul Griffith updated these visions of social accounts of emotion.

What is an Emotion? It may be feelings, or adaptive dispositions, or cognitive processes, or behavior, or social performance. In fact, most of these descriptions of emotion are partial. Research on emotions is as vast and diverse as emotional life itself. This diversity of theoretical perspectives on emotions suggest that emotions are

---

<sup>2</sup> Patterns of action readiness following appraisals were what characterized different emotions

undefined (Lazarus, 1984) and yet this diversity is necessary. Each of these different theoretical approaches provide its own answer to the ontological question of emotion, together define the phenomena, suggest specific mechanisms underlying them, and frame questions that future emotion research should address (Leventhal & Tomarken, 1986). Each tradition for defining emotion helps us understand the ground rules for theorizing emotion.

Philosopher Evans decided wisely to avoid grappling with each tradition and controversies that existed, and provided a comprehensive guide to current scientific thinking about emotion. His account of the definition of emotion provides a solid foundation for better understanding emotions in learning.

### **2.3.2 Evans' accounts of emotion**

From the above discussion of five emotion theories, we can see that researchers divide themselves into at least two camps: one camp sheds light on the common biological bases of a limited set of so-called basic or discrete emotions, such as anger, fear, sadness, joy and love, which appear to be universal in all human beings (Ekman, 1973; Emde, 1980; Izard, 1977). The other camp emphasizes the more complex, socially constructed emotions, which show great cultural diversity (Harre, 1986). Evan's (2001) accounts of this long-time controversial issue, the taxonomy of emotions and comprehensive emotion lists, open a new age of emotion research and provide us with intelligent pointers for researching emotion.

In '*Emotion*' (2001), Evans did not proclaim which camp he belonged to but mapped a picture for emotion research: a spectrum of innateness, with basic emotions located at the very innate end, and culturally specific emotions at the least innate end.

Adding a third category called “higher cognitive emotions” (Evans, 2001, p. 28), he divided the spectrum into three sections rather than two. Higher cognitive emotions are less innate than basic emotions, but more innate than culturally specific ones. When deciding what an emotion is, we should not ask whether an emotion is innate or not, but rather how innate an emotion is. The more ‘special conditions’ over and above the basic necessities of survival that are required for the development of a trait, the less innate it is.

Evans (2001) synthesized the five emotion research theories and pointed out that basic emotions and cultural-specific emotions sat at opposite ends of a single spectrum. He further explained that depending on how many conditions were required for a given emotion to develop and how special they were, an emotion would be located more towards the ‘basic’ end of the spectrum or more towards the ‘cultural-specific’ end. Basic emotions are much more innate than culturally specific emotions, but they still require some minimal conditions to develop (Evans, 2001).

Following American anthropologist Paul Ekman’s well known comparative studies of emotions in Western and non-Western cultures, Evans laid out some emotions that we could find the world over: they are not learned; they are universal and innate. They are what Ekman called ‘Basic Emotions’ (Evans, 2001), which include **joy, distress, anger, fear, surprise, and disgust**. There is no culture in which these emotions are absent. Moreover, they are hardwired into the human brain. Basic emotions are not like words, which differ from culture to culture; they are closer to breathing. Thanks to these basic emotions we share, communication is possible without words (e.g., facial expressions and bodily gestures). Our common emotional heritage binds humanity together, then, in a way that transcends cultural difference. In all places and at all times, human beings have

shared the same basic emotional repertoire. Different cultures have elaborated on this repertoire, exalting different emotions, downgrading others, and embellishing the common feelings with cultural nuances, but these differences are more like those between two interpretations of the same musical work, rather than those between different compositions (Evans, 2001).

Evans' explanation of basic emotions is in accordance with the universality of basic emotions in Darwinian-Evolutionary accounts, both arguing strongly for their biological nature. If basic emotions were cultural inventions, their ubiquity would be very surprising. Like humans sharing the same physiology of body, they also share the same physiology of mind. This universal human biology is encoded in the human genome, the fund of our shared evolutionary history. Our brains are basically the same the world over. Since basic emotions are largely determined by the structure of our brains, it really should come as no surprise that they too are fundamentally the same in all cultures. In a search for cultural identity, Evans (2001) argued that researchers naturally fixed on the things that set us apart from others, rather than on the things that linked us together; when it comes to emotions, researchers often paid attention to small cultural differences, and ignored overwhelming similarities.

Evans employed an example of the 'inscrutable oriental' to argue that every culture had its own rules that defined socially acceptable forms of emotional expression. In Europe and North America, these 'display rules' encourage vivid facial expressions of emotion; a poker face is generally regarded as dull or deceptive. In Asia, excessive emotional displays are often perceived as rude, and people consequently make more of an effort to attenuate their emotional expressions. But underneath these 'display rules', the

emotions are the same. In all, basic emotions are hardwired, etched into our neural circuitry by our genes rather than by our culture, as part of the basic mental design that is common to us all.

On the other hand, Evans demonstrated the reality of cultural-specific emotions. He used the emotion felt by Gururumba people of New Guinea, as an example, to illustrate that the emotion of 'being a wild pig' felt by Gururumba people was not innate. It needs a special condition for it to develop, conditions that are provided only by particular cultures. So this emotion is culturally specific, which unlike basic emotions that develop innately, and develops only if humans are exposed to them by their own culture. It is precisely the function of many emotions that they help people cope with particular demands of their culture. Basic emotions are not tailored to fit the specific demands of a particular culture, but designed to help us meet fundamental challenges faced by humans everywhere. Therefore, Evans believed that the pivotal role that cultural-specific emotions played in human society could not be replaced.

Evans explained that there existed another emotion called higher cognitive emotions. Higher cognitive emotions are universal, but they exhibit more cultural variation. They take longer to build up, and longer to die away, than basic emotions. Higher cognitive emotions include: **love, guilt, shame, embarrassment, pride, envy and jealousy**. Some basic emotions can be co-opted for the social functions that typify higher cognitive emotions.

Higher cognitive emotions involve much more cortical processing than basic emotions, which mean that higher cognitive emotions are more capable of being influenced by conscious thoughts, and this in turn is probably what allows higher



cognitive emotions to be more culturally variable than basic emotions. However despite their greater cultural variability, the higher cognitive emotions are still universal. Like basic emotions, but unlike culturally specific emotions, higher cognitive emotions are part of human nature, shaped by our common evolutionary history. These emotions are fundamentally social in a way that basic emotions are not. Higher cognitive emotions seem to have been designed by natural selection precisely to help our ancestors cope with an increasingly complex social environment. These emotions may be the cement that binds human society together.

In conclusion, emotions are not easily divided into two or three clear-cut categories, but are allocated on a spectrum according to degrees of innateness. Basic emotions and cultural specific emotions sit on both the end of the spectrum, and higher cognitive emotions are in between.

### **2.3.3 State of the art of key research**

Since Rosalind Picard elaborated evidence of the view that computers needed emotions (Picard, 1997), emotional technology research has gained an unprecedented growth. Recent advances have enabled human users to interact with technologies in ways previously unimaginable. However, emotional technology research is still in its infancy. Three contemporary main areas of research and studies – investigating the potential for technology to respond to human emotion (emotion recognition system), stimulating human emotion (emotion generation system), and representing human-like emotion (emotion simulation), are emerging. This part of the review explores emerging technologies in HCI that enable the computer to be more aware of the user's emotional expressions.

### **2.3.3.1 Facial Expression Recognition Studies**

Facial expressions convey non-verbal cues, which play an important role in interpersonal relations. Facial expressions and movements such as a smile or a nod are used either to fulfill a semantic function, to communicate emotions or as conversational cues. The research on facial expression consists of works on coding, automatic recognition and generation of facial expression, which are important components of natural human-computer interfaces. Ekman's facial expression and the Facial Action Coding System (FACS) to code facial expressions where movements on the face are described by a set of action units (Aus) inspired works in computer-assisted quantification of facial expressions. Automatic systems for facial expression recognition usually take the form of a sequential configuration of processing blocks, which adhere to a classical pattern recognition model (Jain, Duin, & Mao, 2000; Pantic & Rothkrantz, 2000). The main blocks are image acquisition, pre-processing, feature extraction, classification, and post-processing.

### **2.3.3.2 Speech Processing Studies**

The vocal aspect of a communicative message carries various kinds of information. Starting in the 1930s, studies of vocal emotions have had a longer history than studies of facial expressions. It is widely known that emotional speech differs with respect to the acoustic features. "If we disregard the manner in which the message was spoken and consider the verbal part only, we might miss the important aspects of the pertinent utterance and we might even completely misunderstand what was the meaning of the message" (Sebe, Cohen, & Huang, 2004, p. 10). Most recent studies in emotional speech

have used the prosody feature, which includes the pitch, duration, and intensity of the utterance. Tato made some experiments, which showed how quality features are used in addition to prosody features (Tato, Santos, Kompe, & Pardo, 2002). Williams and Stevens studied the spectrograms of real emotional speech and compared them with acted speech (Williams & Stevens, 1972). Murray and Arnott reviewed findings on human vocal emotions and constructed a synthesis-by-rule system to incorporate emotions in synthetic speech (Murray & Arnott, 1993). To date, most work has concentrated on the analysis of human vocal emotions. “Nevertheless, in contrast to spoken language processing, which has recently witnessed significant advances, the processing of emotional speech has not been widely explored” (Sebe et al., 2004, p. 10).

#### **2.3.3.3 Body Gestures and Movements**

Body gestures and movements are the positions of body arthroses and their changes with time. The recent studies on gesture and movement processing are more focused on the hand tracking. Hand gestures can convey various and diverse meanings and can enhance the mood as well as serve as a sign language. Traditionally, there are two methods, apparentness methods and 3D modeling methods, for the study of body gesture and movement recognition. However, the capture of body gesture and movements is still a difficult subject in the area of computer vision, especially in real application, let alone the challenge to capture accuracy and efficiency of parameters to obtain more robust and subtle body-language.

#### **2.3.4 State of the art of key projects**

Many emotion technologies under ‘affective computing’ have been designed. MIT’s Media Lab has done a substantial amount of research on techniques and devices

for assessing users' emotional states. Vyzas and Picard (1999) showed how physiological data on jaw clenching, blood volume pressure, skin conductance and respiration could quite accurately be recognized as eight different emotional states, when a person intentionally expresses them. Healy and Picard (2000) used input from electromyogram, electrocardiogram, respiration and skin conductance sensors to detect stress in a car driver. Kaapor, Mota and Picard (2001) discussed how to monitor eyebrow movements and posture to provide evidence of students' engagement while they interacted with a computer-based tutor. Other research organizations and research communities shed light on affective computing ideas and techniques. Bot technology attempts to replicate human communication; Sony's Aibo creates emotional reactions in users; automatic speech recognition technologies respond to verbal inputs and act depending on users' emotional states; affective reasoner is a rule-based framework to build agents that respond emotionally (Elliott, Rickel, & Lester, 1999); a probabilistic model can be used to monitor users' emotions and engagement during interaction with educational games (Conati & Zhou, 2002); an intelligent desktop assistants uses a decision theoretic approach to decide when and how to provide unsolicited help to the user (Horvitz, 1999a, 1999b).

In all, designing technology to address emotions is a challenging endeavor. Firstly, technology that can display emotions in a natural and meaningful way requires formalizing concepts and mechanisms that are still under investigation in emotional psychology (Conati & Zhou, 2002); secondly, technologies that can recognize user's emotions requires a high level of detecting human's emotions. In fact, emotion itself is highly complex, and can be recognized by technology only because it often has

observable effects on a user's behaviors and bodily expressions. But due to the ambiguity that exists between emotional states and factors such as age, gender, or culture, which influence emotions a lot, even humans cannot detect each other's emotions with perfect accuracy. Therefore the idea that technology might one day come to readily address emotions will guide future researchers in the field.

### **2.3.5 Skepticism towards Affective Computing**

Although affective computing receives high prestige these days, some researchers keep their skepticism towards this viewpoint. One of these opponents is Erik Hollnagel. He argues that "affective computing by its very nature cannot be affective and affective computing should refer to a specific use of computing rather than a specific type of computing" (Hollnagel, 2003, p. 65). Hollnagel discussed the controversy between the meaning of computer<sup>1</sup> and computing<sup>3</sup> and emotions. For him, since computing works with data in discrete forms and a bit hardly can have emotions as well as logic, which itself is supposed to represent cold cognition rather than emotion (Abelson, 1963; Hollnagel, 2003), affective computing contradicts itself in terms. Hollnagel thought the belief that if humans can have emotions although they are cognitive beings, this is not possible for computers due to "a misunderstanding of cognition and a misunderstanding of emotions or affects" (Hollnagel, 2003, p. 66). He explained the reason that AI researchers and cognitive science had this belief is that they accounted emotion from a cognitive perspective. Hollnagel offered another term—effectual communication—to

---

<sup>1</sup> A computer can be defined as any of various automatic electronic devices that solve problems by processing data according to a prescribed sequence of instructions.

<sup>3</sup> Computing, as data processing or information processing, is thus the execution of an unambiguous ordered sequence of computational instructions that represent a detailed plan or procedure for solving a specific problem.

describe human–computer interaction that includes something akin to affect. The purpose of computing, rather than transmitting emotions, is to “adjust the style of communication to achieve maximum effectiveness” (Hollnagel, 2003, p. 69).

Lindgaard also pointed out his struggle to accept the notion that his computer should take his emotional temperature regularly and modify its canned responses according to its reading of his emotional thermometer (Lindgaard, 2004). He tackled the complex notion of affect and trust, and emphasized that researchers need to carefully assess and clarify what they mean by human needs in the context of affective computing and questioned that “ how might affect be measured, and how are measures to be interpreted” (Lindgaard, 2004, p. 727)? At the end he concluded that he remained skeptical about the long-term success of canned-response affective computers (Lindgaard, 2004).

## **2.4 Multimodality**

Humans use language to communicate not only by using the words of the language but also by using several modes of production, such as hand gestures and facial expressions, head movement, eye contact as well as body posture etc. More importantly, when communicating, several modes of human perceptions are also involved such as nodding to show understanding, shaking heads to show frustration or puzzlement and facial emotional expressions to signal how one understands the other, etc. Humans easily and naturally use all of their sensory modalities as they communicate and exchange information.

As advances are made in learning theories and technologies for human–computer interaction, CALL applications have to go beyond the simple reflexive behavior to

human-like and authentic communication. Hence, the design of them has become more and more complex. A single modality does not permit the user to interact effectively across all tasks and environments (Larson et al., 1999). Also, language learning is a very “wide” subject where some certain issues such as grammar rule drills or pronunciation exercises can rely on traditional interaction techniques and are suitable to implement with unimodal interface, while other issues such as language understanding and communicative abilities are more likely to run into difficulties when using a unimodal interface and thus a multimodal interface has potential opportunities. Therefore, it has become increasingly important to put requirements of a multimodal interface, which combines a speech interface with the traditional display interface, on a CALL application that aims at teaching the use of language. The flexibility of such a multimodal interface for a CALL application can accommodate a wide range of users, tasks, and environments for which any given single mode may not suffice (Oviatt, 1999).

A multimodal interface provides the user with multiple modes of interacting with a system beyond the traditional keyboard and mouse input/output. It combines a visual modality (e.g. a display, keyboard, and mouse) with a voice modality (speech recognition for input, speech synthesis and recorded audio for output). Other modalities such as pen-based input or haptic input/output may be used. “The growing interest in multimodal interface design is inspired largely by the goal of supporting more transparent, flexible, efficient, and powerfully expressive means of human-computer interaction” (Oviatt et al., 2000, p. 265).

## **2.5 Emotions in Learning**

In an attempt to design a multimodal adaptive CALL interface, we should not only discuss the benefits of multimodality and adaptivity, but also we need to rethink what is happening during learning. Findings in neuroscience, psychology, and cognitive science show that emotions are intertwined with thinking and “performing important functions with respect to guiding rational behavior, memory retrieval, decision-making, creativity and more” (Picard et al., 2004, p. 253). Since neurologist Damasio argued that emotion played a critical role in cognition, which traditional approaches to cognitive interpretations of human behavior neglect, the past decade has witnessed a new scientific appreciation of emotion in its service of cognition. However, “most of classic works on affect emphasized cognitive and information processing aspects in a way that can be encoded into machine-based rules, and studied in a learning interaction” (Picard et al., 2004, p. 255). “There is very little understanding as to which emotions are most important in learning, and how they influence learning. To date there is no comprehensive, empirically validated, theory of emotion that address learning” (Picard et al., 2004, p. 255).

A MIT research group conducted a study with elementary school children and described the range of various emotional states during learning (see Table 2.1). They then proposed a model of a learning cycle (see Table 2.2), in which circulates the flow of emotions (see Table 2.3) (Kort, Reilly, & Picard, 2001b).



Table 2.1 : Axes of emotional sets (Adapted from Kort, Keilly, & Picard, 2001a)

Axis	<div style="display: flex; justify-content: space-between; align-items: center;"> <span>Negative</span> <span>Positive</span> </div>					
Anxiety- Confidence	Anxiety	Worry	Discomfort	Comfort	Hopeful	Confident
Boredom- Fascination	Ennui	Boredom	Indifference	Interest	Curiosity	Intrigue
Frustration- Euphoria	Frustration	Puzzlement	Confusion	Insight	Enlightenment	Epiphany
Dispirited- Encouraged	Dispirited	Disappointed	Dissatisfied	Satisfied	Thrilled	Enthusiastic
Terror- Enchantment	Terror	Dread	Apprehension	Calm	Anticipatory	Excited

Table 2.2: The model of emotion in learning (Adapted from Kort, Keilly, & Picard, 2001a)

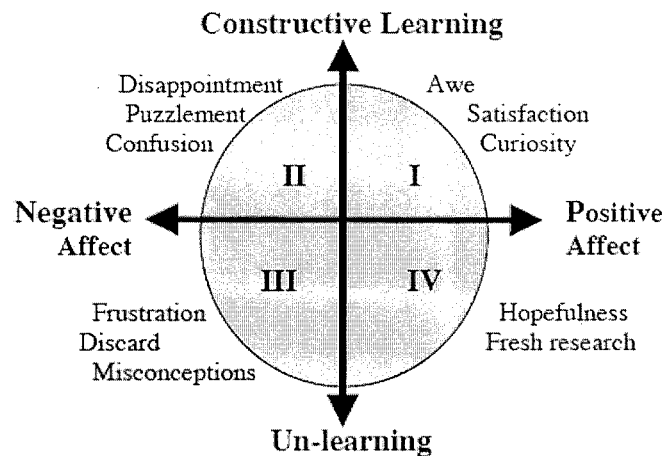
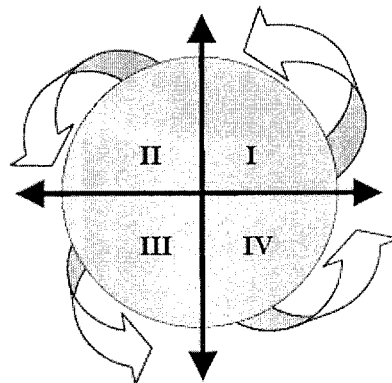


Table 2.3: Circular and helical flow of emotion (Adapted from Kort, Keilly, & Picard, 2001a)



### 2.5.1 The model of emotion in learning

The model of a learning cycle (Table 2.2) proposed by Kort, Reilly and Picard (2001a) attempted to assume that emotion was related to various phases of learning. They firstly identified several emotional axes, specifying a range of emotional states (Table 2.1). These emotional axes interweave with the cognitive dynamics of learning processes (Table 2.1). The horizontal axis is an emotion axis, which can be one of the specific axes from Table 2.1. The positive emotions are on the right; the negative emotions are on the left (Kort et al., 2001a). The vertical axis is the learning axis. The learning axis<sup>4</sup> ranges from ‘constructive learning’ at the top, where new information is being integrated into schemas, and ‘un-learning’ at the bottom, where misconceptions are identified and removed from schemas (Kort et al., 2001a).

When elaborating the relationship between emotion and learning, Kort, Keilly & Picard used a four – quadrants learning spiral model in which a typical learning experience involved a range of emotions, moving students around the space as they learn (Table 2.2 & Table 2.3). Thereby, the challenges for teachers are to “help them (students) move from anxiety to confidence, from boredom to fascination, from frustration to euphoria, disillusionment to encouragement, and from terror to enchantment” (Petrina, 2007, p. 68).

Movement ideally begins in quadrant I or quadrant II and proceeds in a counter-clockwise direction (Kort et al., 2001a). In either quadrant I or quadrant II, the learner’s focus is on constructing knowledge. However, they may be curious and fascinated about

---

<sup>4</sup> This terminology is merely a projection of one aspect of how people can think about learning. It doesn’t mean simply a process of constructing/deconstructing or adding/subtracting information.

a new topic (Quadrant I) or they may be puzzled and motivated to reduce confusion (Quadrant II). When they feel curious and fascinated, students are working through the material with ease and do not experience anything overly puzzling. Once discrepancies start to arise between the information and learners' schemas, they move to quadrant II, where they experience affective states such as puzzle and confusion. As learning proceeds, movement happens. When students encounter problems, which need to be deconstructed they move down into the lower half of the diagram (Quadrant III) where emotions may be negative and the cognitive focus changes to eliminating some misconception. As learners start to discard their misconceptions and consolidate their knowledge – what works and what does not work – with awareness of a sense of making progress, they move into quadrant IV, which is marked by unlearning and positive affect. While in this quadrant, learners are still not sure exactly how to go forward. However, they do acquire new insights, search for new ideas and have eureka experience when the insights are profound. Once they develop fresh ideas, they are propelled back into the upper half of the space: quadrant I. This will conclude one learning circle. As learners move up the spiral, cycle after cycle, they become more competent and acquire more domain knowledge (Kort et al., 2001a, 2001b).

### **2.5.2 Emotions in the service of cognition**

The Western approach to thinking and feeling, cognition and emotion has been to regard them as hostile worlds apart. But Evans' functionalist view considers emotions interweave with cognition and serve its processes. "Cognition cannot be cleanly sundered from emotion and assigned to science, while emotion is ceded to the arts, ethics, and religion. All these spheres of life involve both fact and feeling; they relate to sense as

well as sensibility” (Scheffler, 1977, p. 178). Emotion without cognition is blind and cognition without emotion is vacuous. Learners’ learning experiences are shaped by emotional, cognitive and social processes working together. Emotion serves cognition in “rational passion, perceptive feelings and theoretical imagination” (Scheffler, 1977). At least, some emotions are essentially cognitive in origin and may in fact serve cognitive purposes. While reason and feeling can be distinguished in theory, they are intimately related in human’s experience of them (Scheffler, 1977).

#### **2.5.2.1 Rational Passions**

Rational passions is a concept from Peters (1970, p. 143):

My interest, needless to say, is in the emergence of a rational form of morality, which enables a person to adopt a stance that is critical of tradition but not subjective. But this cannot be characterized purely in terms of the ability to reason, in the sense of making inferences, as I have argued elsewhere. To start with, if this ability is to be effectively exercised, it must be supported by a group of rational passions connected with the demands for consistency, order, clarity and relevance.

Rational Passions refer to “the emotions under-girding the life of reason” (Scheffler, 1977, p. 173). They are those emotional dispositions, which make up ‘an intellectual conscience’. The passions, according to Scheffler, are “a love of truth and a contempt of lying, a concern for accuracy in observation and inference, and a corresponding repugnance of error in logic or fact, revulsion at distortion, disgust at evasion, admiration of theoretical achievement, respect for the considered arguments of others” (Scheffler, 1977, p. 173). They are internalized as rational norms and are characteristic of cognitive discipline. “These emotions are indispensable to the integrity of the activities of reason, they are structured intelligently, and to be operative; they are internalized and

personalized” (Yob, 1997, p. 45). Certain impulses such as “inconsistency, unfairness to the facts, and wishful thinking” are not controlled “through a bloodless reason, as control is exercised through the structuring of emotions themselves” (Scheffler, 1977, p. 173). “Emotion of some kind accompanies reason from the beginning and is not merely something towards, which reasoned principle moves and further, that both reason and emotion become more disciplined throughout the course of their mutual development” (Yob, 1997, p. 45).

#### **2.5.2.2 Perceptive Feelings**

Perceptive feelings, distinguished from the internalization of rational norms, are a consideration of emotion employment in perception, which in effect are an elaboration of Peters’ “appraisals.” Feelings “are intimately tied to our vision of the external world... [and] help to construct that vision and to define the critical features of that world” (Scheffler, 1977, p. 174). “Feelings are the emotional filters through which we view the world, interpret its objects, and evaluate its critical features” (Yob, 1997, p. 45). These critical features are “the objects of our evaluative attitudes, the foci of our appraisals of the environment” (Scheffler, 1977, p. 174).

Appraisals refer to assessments of relations between perceived events and a person’s goals, motives, and concerns (Arnold, 1960; Frijda, 1986; Lazarus, 1991). In emotional appraisal, cognitive processes function in the service of an individual’s goals, motives, and concerns (Smith & Lazarus, 1990). They are essential in representing one’s motives and concerns and in mediating an individual’s interpretation of any given event. Secondly, appraisals are products of both conscious and non-conscious processes. Different categories of emotional experiences attribute to different ways in which people

appraise events (Lazarus, 1999; Mascolo & Griffin, 1998; Mikula, Scherer, & Athenstaedt, 1998; Parkinson, 1999; Roseman, Antoniou, & Jose, 1996; Roseman, Spindel, & Jose, 1990). Emotion is an intentional state (Brentano, 1874; Campos, 1994; Searle, 1980; Solomon, 1976), hence, an emotion has objects; it is about something. It involves “seeing the environment in a certain light... beneficial or harmful, promising or threatening, fulfilling or thwarting” (Scheffler, 1977, p. 174).

Scheffler used ‘fear of a particular person’ as an example to indicate that through emotion, “we gain enormous new powers of fundamental description” (Scheffler, 1977, p. 174). According to Scheffler, fear of a particular person is aroused due to the presupposition that the person is regarded as dangerous. Thus, danger becomes a critical feature of the environment calling for a special orientation in response. Hence, prompted by cues from danger in the environment, we may, further, come to describe a certain situation as terrifying and arouse the feeling of fear (Scheffler, 1977). Scheffler attributed these emotional cues to “the product of evolutionary development” and “the special circumstances of individual biography” (Scheffler, 1977, p. 174). “Our powers of discernment are enhanced if we are skilled in appropriating our feelings as well as our reasoning, making sense of the world and its objects” (Yob, 1997, p. 45). The cognitive role of emotions is not limited to the aesthetic realm where their function has been recognized but emotions intimately mesh with all critical appraisals of the environment: the flow of feelings provides us with a continuous stream of cues significant for orientation to our changing contexts (Scheffler, 1977).

### **2.5.2.3 Theoretical Imagination**

Emotions stimulate scientific imagination (Scheffler, 1977). The process of theory building is not merely a matter of fact-gathering and methodological application of scientific procedures, but is a matter of boldness and speculative daring throughout (Scheffler, 1977). “Imagination is no hindrance but the very life of theory, without which there is no science” (Scheffler, 1977, p. 177). In fact, without “feeling and flair,” theorizing is not only deadly, but indeed “dead” (Scheffler, 1977; Yob, 1997, p. 177). Emotions relate to imaginative theorizing in three ways: the emotional life is a rich source of substantive ideas; emotions fulfill a selective function, facilitate choice, define salient features and focus attention accordingly; emotions played a directive role in applying imagined solutions to problems encountered (Scheffler, 1977).

In all, the interplay between emotion and cognition show that “Knowledge does not exist independent of feelings or physical skills. Our emotions and skills cannot be separated from our capacity to learn and to act thoughtfully” (Petrina, 2007, p. 19). As Piaget noted, “at no level, at no state, even in the adult, can we find a behavior or a state which is purely cognitive without affect nor a purely affective state without a cognitive element involved” (Clark & Fiske, 1982, p.130). There is a need to understand humans holistically (MaKeachie, 1976): cognition and affect should not be seen as inseparable.

## **2.6 Affective Domain of Learning**

Besides adaptive multimodal features of the CALL application and its learning model embedded, teaching objectives are necessary, which provide direction for users to assess whether or not they are learning what is designed. There are three instructional domains, which teachers address during their instruction. “The cognitive domain refers to

the recall or recognition of knowledge and intellectual abilities and skills. The affective domain refers to changes in appreciations, attitudes, emotions, interests, and values. The psychomotor domain refers to the development of manipulative, sensory and motor skills” (Petrina, 2007, p. 21). However, I shed light only on the affective domain in this study due to the emphasis on emotional aspects of learning.

Developed by Krathwohl et al. (1964), the affective domain is based upon behavioral aspects: “This taxonomy is ordered according to the principle of internalization. Internalization refers to the process whereby a person’s affect toward an object passes from a general awareness level to a point where the affect is ‘internalized’ and consistently guides or controls the person’s behavior” (Seels & Glasgow, 1990, p. 28). This domain includes the manner in which we deal with things emotionally, such as feelings, values, appreciation, enthusiasm, motivations, and attitudes. The five major categories are listed from the simplest behavior to the most complex (Table 2.4).

Table 2.4: Affective domain: Knowing (Adapted from Krathwohl, Bloom, & Masia, 1964)

Receiving	Attention to particular phenomena or stimuli (activities, textbook, music, etc.). Attention ranges from simple awareness to selective attention.
Responding	Active participation that involves attention (receiving) and reaction. Acquiescence in responding, willing attitude, and display of satisfaction or dissatisfaction. Interest is exhibited.
Valuing	Worth or value attached to objects, people or processes. Ranges from acceptance of value to complex levels of emotional commitment and responsibility toward values. Valuing is based on the internalization of a set of specific values and the actualization of these values in overt behavior. Behavior and emotions are consistent with values.



Organization	Convergence of different values, resolution of value conflicts, and internally consistent value system. Emphasis on comparing, relating and synthesizing values. Individual is able to articulate how emotions and values are conceptualized and organized into value systems.
Characterization	Individual has articulated a value system that has informed actions and emotions for periods sufficient to the development of a lifestyle. Behavior is consistent, value-driven, pervasive and predictable. Emotional patterns are mature and reflective. Individual is in touch with feelings.

## 2.7 Conclusion

With a review of current literature, this chapter argued that the adaptivity, multimodality, learning model and learning objectives were crucial to the study of user – centered CALL application interface design. The theoretical starting point was from Evans’ comprehensive accounts of emotions to Scheffler’s cognitive emotions, which provided a useful and promising theoretical framework for my inquiry into users’ (learners’) emotions in its service of their cognition. The review of chatbot technology and its usage, adaptive and multimodal interfaces as well as Kort’s learning model and Krathwohl’s affective domain allowed me further unravel the complex interplay between emotion, cognition and technology to achieve a better understanding of user-centered CALL interface design factors.

## Chapter Three: Research Design

My study drew upon the commercial chatbot Lucy, a digital language tutor, which was developed with the hope to provide language learners an innovative way to learn language. What fascinated me about this chatbot technology was its multimodal interface, which combined speech recognition with graphic design. This study contributed to the user-centered speech-driven multimodal adaptive language learning interface design, aiming at investigating learners' experiences when using Lucy and finding out which features of chatbot Lucy could make learning effective. An ethnographic research methodology was employed in order to empirically analyze learners' experiences with Lucy, which was the focus of the interface design.

The theoretical framework that I used to understand users' learning experiences precipitated a quest to use ethnography to document the human-computer interaction details. Therefore, consistent with the theoretical framework described in Chapter Two, the research design of this study, whose purpose is to provide a holistic description of users' learning experiences with the chatbot, employed an ethnographic methodology<sup>5</sup> and ethnographic research methods in the process of data collection, namely, participant observation, participant reconstructive interviews and video recording. By using these methods, this study generated an integrated data set, which included not only new possibilities to understand users' learning experiences but also fresh sight of interface design.

---

<sup>5</sup> The distinction between *methodology* and *method*, as indicated in Creswell (1998), Stake (2000), and Wolcott (1999), is that *methodology* involves the whole process of research, while a *method* is only an approach or technique to collect data.

To achieve the aims stated above, this chapter considered: 1) the rationale for choosing particular research methods; 2) unit of analysis; 3) research site; 4) instruments for the study; 5) participant ethics, and 6) an introduction to data.

### **3.1 Methods**

Ethnography as a means of describing users' learning experiences was employed in this study. Conventional ethnographic research gets its rich traditions within sociology and anthropology (Bogdan & Biklen, 1992; LeCompte, 1993; Lincoln & Guba, 1985) and typically includes: field work done in natural settings, the study of a large picture to provide a more complete context of activity, an objective perspective with rich descriptions of people, environments and interactions, and a bias toward understanding activities from the informants' perspective (Blomberg, Giacomi, Mosher, & Sewenton, 1993; Millen, 2000). But the meaning of ethnography has been extended and practiced in many disciplines other than sociology and anthropology. Nowadays, many ethnographers categorize any small-scale study carried out in everyday setting that focuses on the individual's meaning and qualifications as ethnography (Savage, 2000). In the field of human-computer interaction (HCI), the use of ethnographic approaches has been focused to understand work practice in order to inform the design of information systems through an ethnographic account. However, "typically, ethnography will take place over a period of several months with at least the same amount of time spent in analysis and interpretations of the observations" (Bentley et al., 1992). Thus, one of the biggest challenges facing HCI ethnographers is the time-consuming fieldwork activities when "matching the pace of ever-quickenening product development cycles" (Millen, 2000, p. 280). This is also the biggest challenge in this research. The results from the fieldwork of

this study will be passed on to the next phase – user-centered redesign stage. Tackling this problem, I adopted Millen’s rapid ethnography approach that aimed to provide a reasonable understanding of users and their activities within a given limited time. I mostly employed participant observation, participant reconstructive interviews (open-ended), and semi-structured pre-research interviews during my fieldwork. This ethnographic study involved fairly lengthy contact with the participants, through participant observation in the lab and through relatively semi-structured interviews and participant reconstructive interviews designed to understand users’ perspectives. I “typically insist(ed) on the importance of coming to understand the perspectives of the people being studied” (Hammersley, 2006, p. 4). A research diary was also kept for the study.

### **3.1.1 Pre-research Interview (Semi-structured)**

To achieve an initial sense of my study, I began with a pre-research interview. Semi-structured interviews were used in the first contact between participants and myself. In asking participants their cultural backgrounds and learning history, I was interested in discovering how participants’ past learning experiences influenced their using of CALL applications, and what cultural elements in these learning processes they found salient.

Each interview ran about 30 minutes in length and was recorded on an mp3 recorder. The pre-research interview prior to the actual data collection was less to collect data than to select appropriate participants for this study, to know of participants’ backgrounds and as well as moreover, to allow the participants and me to get familiar with each other.

### 3.1.2 Participant Reconstructive Interviews

In this study, participant reconstructive interviews, which depended on face-to-face questioning of participants and eliciting data from them, were employed after their interaction with the five modules of chatbot Lucy. An interview is “the ethnographer’s most important data gathering technique” (Fetterman, 1998, p. 37), because “it directly taps into the participants’ perceptions and views of reality” (Zhang, 2007). Every participant on video camera was interviewed for about 60 minutes as soon as possible after his/her interaction with five modules. By using the video record as stimulus for participants’ reconstructions of their learning experiences, participant reconstructive interviews aimed to record as much as possible of their situated learning when using chatbot Lucy. During the interview, I, as an interviewer, had a monitor with a DVD player. A video camera was set up to record interview data, mainly focused on the interviewees. Participants had a controller to control the DVD player and were asked to identify any sections shown by the video file that involved or heightened their emotional responses. The video file was played at a normal speed, or pause. Participants were invited to give detailed description of

1. How they feel about the learning interface of chatbot Lucy, especially its weakness.
2. What kind of emotions are involved during interactions with chatbot Lucy?
3. Which features/functions of the interface triggered these emotions?
4. How their feelings with chatbot Lucy influenced their learning processes.
5. Whether they think the chatbot technology helps them learn or actually hinders their progress of learning.

6. What are participants' learning goals of using chatbot-assisted language learning interface?

Participants were given control of the video play and were asked to identify and comment upon their emotional ties with Lucy. Such verbal accounts provide legitimate data for this research (Ericsson & Simon, 1980; Nisbett & Wilson, 1977). Within the study, participants' verbal reconstructions of their motivations, feelings and thoughts were given significant thought.

Videotaping participants' learning processes provided specific and immediate salient stimuli that could optimize the conditions for effective recall of associated feelings and thoughts of their interacting with Lucy, and participants' verbal reports of their thoughts and feelings, when prompted by videos of the particular associated scenario, could provide useful insights into these participants' learning experiences. The important key role of the video-stimulated recall resided in the juxtaposition of participants' account of their appraisal, feelings and actions.

The aim of these participant reconstructive interviews was to collect data, which illuminated the range of emotions involved in learning processes in HCI settings and tried to discover their learning experiences with Lucy.

### **3.1.3 Participant Observation**

The above two interviews captured participants' learning experiences that they themselves found memorable or remarkable. However, to see the actual moment-to-moment learning processes during the interaction with Lucy required participant observational studies of the learning itself. My use of participant observation method reflected the conviction that there were essential details of everyday, situated human

activities and interactions that would always be missing from interview data (Suchman & Trigg, 1986). The meanings of action are grounded in the context of activity (Hollan, Hutchins, & Kirsh, 1999). The method of participant observation, with its virtue of direct experiential and observational access to the meanings and interactions in everyday life (Jorgensen, 1989), offered the opportunity for me to understand and capture moment-to-moment participants' learning experiences and participants' emotional responses to chatbot Lucy's learning interface. In addition, direct observation also enabled me to see things, which were often overlooked, and learn things that participants were unwilling to discuss in the interview (Patton, 2002). To that end, participant observation provided an overview of progress in learning during the interaction with Lucy.

#### **3.1.4 Videotaping**

Data collection processes employed two cameras: one in front of the participant and the other behind him/her. The two participants' camcorders maintained a continuous record of the participant's facial expressions, gestures, statements and actions. Each participant was videotaped individually during interaction with Lucy (each session lasted 30-40 min). Each camcorder was set up prior to the commencement of the interaction to include the participant and adequately focus on him/her. The camcorder remained fixed unless participant's movement necessitated its realignment. During participants' interaction with chatbot Lucy, I manipulated the camcorder on each participant and zoomed in each camera on each participant's written work every twenty minutes or so to maintain an on-going record of the participants' progress on any written tasks. This zooming in was done briefly to provide visual cues of the progress of the participant's written work, but was done without losing the continuity of the video record of each

participant's facial expressions, gestures, statements and actions that were needed for the subsequent interviews. Since the participant's learning processes are the priority in this study, the continuous documentation of the actions and expressions of each participant and his/her interaction with chatbot Lucy were the most important. The video record served to display each participant's facial expressions, gestures, activities, actions and communications in relation to the interaction with the chatbot.

During data collection, as a researcher, I sat on site behind or beside each observed participant. A running sheet was developed for myself to jot down any moments that were thought as significant, which would subsequently be used as reference points in the follow-up participant interviews.

### ***3.2 Unit of Analysis: The Participants***

According to Wolcott (1999), ethnography can be used as a research technique to observe activities in the field, conduct interviews, examine written documents, and even to view, analyze, and interpret the social group(s) within the unit of analysis as a culture. The unit of analysis was every user's experiences with Lucy. These users, mainly from China and Thailand, became the chatbot using cultural-sharing group in this study, in which they shared the culture of using chatbot technology for their language learning.

Fifteen participants at the intermediate level of English with a desire to practice their English with chatbot technology were selected for this study. A demographically mixed selection of these fifteen participants provided a diversity of participant backgrounds. The fifteen participants came from a university in British Columbia and were selected entirely by volunteering. Participant consent was gathered before the research.



At the starting point of selection, I put up posters at universities and various institutions advertising chatbot Lucy as a means for languages learning. And then, I employed a Network or Snowball selection method in choosing participants for this study. This method allowed me to choose my participants based on the referrals of a preceding individual who participated in the research project. I chose this method of selection because individuals who were likely to be interested in chatbot technology in language learning were scattered throughout populations, colleges, universities or institutions.

The employment of these fifteen participants were based on the criteria that (1) participants may have language learning experiences both in classroom settings and in HCI contexts; (2) participants are not native English speakers; (3) Participants are available to participate in the research from September, 2007 to March, 2008.

### **3.3 *Site of Research***

This study was conducted in a lab in the education building at the university. The lab had PCs, Linux systems, and Mac computers. Five modules of chatbot Lucy were installed on two PC computers shown in Table 3.1. It was intended that each participant needed to interact with all five modules. This produced at least 75 consecutive interactions.

Two cameras were prepared for videotaping each participant's interaction with Lucy. A running sheet was developed and put on the table for taking field notes, on which I recorded the time and types of all changes during the interaction. Such field notes were very simple. The completion of field notes was a much lower priority than the maintenance of a continuous video record of participants and did not distract myself from

the primary purpose. Participants' voices were recorded through the camera's internal microphone.

Table 3.1: Installation of Chatbot Lucy

Module \ PC	PC1	PC2
1: Small Talk	+	–
2: Hotel English	–	+
3: Travel English	–	+
4: Restaurant English	–	+
5: Helping Visitor	–	+

### 3.4 *Instrument for the Study*

#### 3.4.1 Chatbot Lucy

The commercial chatbot Lucy is a digital language tutor that carries on extensive conversations with users as they speak into their computers through a microphone (see Figure 1).



Figure 1: Lucy's five learning modules

Using an advanced speech recognition system, Lucy can give users feedback on their pronunciation and guide them through useful exercises to improve their pronunciation and accuracy. Lucy's world is where users meet Lucy. In each Lucy's world module, Lucy offers users over 1000 sentences on a specific subject. Each module focuses on a different topic including helping visitors, hotel English, giving directions-English for traveling and restaurant English (see Figure 2).

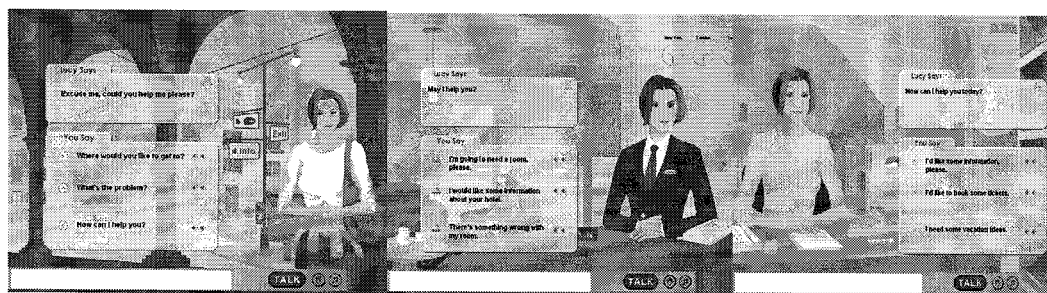


Figure 2: Samples of Lucy's topics

All users' need is a microphone connected to a PC. Lucy's learning materials are translated into seven languages including simplified Chinese, traditional Chinese, Japanese, Korean, Vietnamese, Russian, Spanish and Portuguese (see Figure 3). When users enter into Lucy's world, Lucy greets them and starts the conversation. If users do not hear or understand what she says the first time, they can click on her to make her repeat. If there are some words that users don't understand, they can just mouse over and Lucy shows the translated languages. If users want to challenge their listening skills, they can close the translation window so they can't read what she says, and they can also reopen this window later.

Five modules of Lucy were employed in this study: Lucy's world: travel English; Lucy's world: helping visitors; Lucy's world: restaurant English; Lucy's world: Hotel English; Lucy's world: small talk. Fifteen participants were allowed to freely interact

with Lucy as many times as they liked in the lab. Lucy's characters are described below:

(Table 3.2)

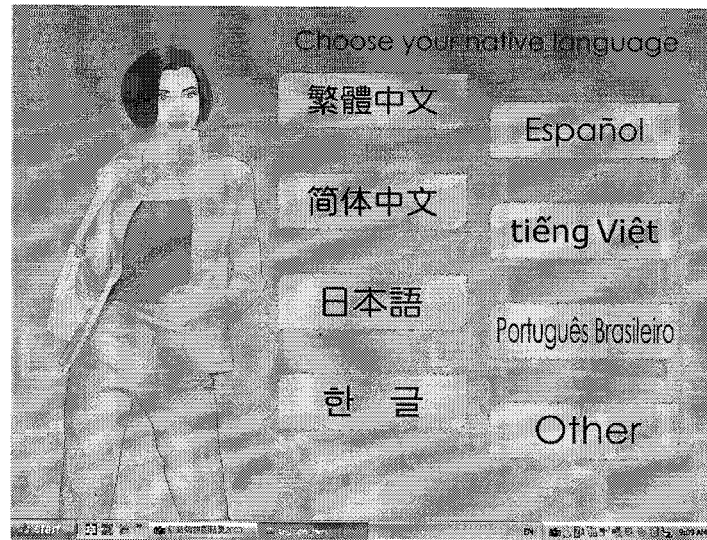


Figure 3: Lucy's translated languages

Table 3.2: Chatbot Lucy's Characters

	<b>Lucy</b>
<b>Visual appearance</b>	
Graphics	2D
Animation	+
Embodiment	Upper half of body
<b>Types of communication</b>	
Input	Speech
Output	Speech
Nonverbal communication	Facial expressions, gestures & body movement
<b>Agent characteristics</b>	
Goal-oriented	+
<b>Intelligence</b>	
Knowledge	Special
Learning	Language
Humor	-
<b>Emotional intelligence</b>	
Emotions	Limited
Personality	-
<b>Functionality</b>	
	Language Tutor

### **3.4.2 Cameras**

Video cameras were the main media for collecting data in this study. Cameras were not only employed during the observation but also used in the participant reconstructive interviews. There were four reasons to choose video cameras in collecting data in this study: videotape could preserve more aspects of interaction including facial expressions, gestures, talking, statements and activities of participants; video allowed repeated observation of the same interaction and supported microanalysis; video triggered participants' recall of their learning experiences during the interviews and video provided analytical benefits. The main features of cameras used in this study were:

- The camcorder for this study needed to superimpose a date and time stamp on the tape;
- The camcorder must have microphone input and headphone output;
- A high-resolution camcorder (Hi8 or SVHS) was used in this study.

### **3.4.3 Microphones**

The external microphone, snowball, located near each user, was required for users' communication with Lucy in this study. Without good sound, Lucy could not recognize the pronunciation or the sentence that participants spoke. Since built-in computer microphones can pick up sound and noise from all the directions, Lucy can't distinguish participants' voice from other noises.

### **3.4.4 Tripod**

A video tripod was used in the research. It had a 'fluid head' for smooth rotation. Two camcorders in this study were fixed on two tripods. When choosing the tripod, I thought about each tripod's height and stability for the research.

### **3.4.5 Monitors**

This study required the users to look at videos on a separate monitor instead of using a computer screen for better quality. A LCD monitor was used in this research during the interview.

## **3.5 *Participant Ethics***

Although this research already met the requirements of human research ethics at UBC (BREB), the direct personal involvement in investigating learning experiences raises some ethical issues. First of all, with videotaping, confidentiality was one of the most salient ethical issues. Maintaining confidentiality by not revealing the individual identities of those studied was a means by which I used to reduce the risk of harm to my participants. When relating to other people about this study, I did not mention any names related to my participants. Research reports or presentations related to this work will use pseudonyms to disguise the real names of my participants.

I was also particularly concerned with the participants' time committed to this study and took care to keep possible intrusion to a minimum of time. In addition, during interviewing my participants, they were in control of the process of the video playing, which helped to reduce my overpowering presence.

In all, this study met all BREB requirements regarding negotiating access to consent of participants to participate. Written consent forms obtained from participants covered data collection, video viewing and analysis of data by researchers and use of the data at conference presentations. The invitation to participate this study involved the following:

- Provision of the usual information sheet and consent to participate sheets to participants
- Consent form, which includes a ‘Consent to be videotaped’ and ‘Consent to be interviewed’ section, each requiring signature by participants
- Selection of participants to be videotaped and preparation of a videotaping schedule

### **3.6 *An Introduction to Data***

#### **3.6.1 Data collection**

The period of data collection lasted six months. As mentioned above, multiple techniques were employed to collect a range of data, which included:

- Participant observation with field notes
- Observation with videotaping
- Documentation of participants’ work
- Video-taped participant reconstructive interviews

These techniques of data collection and analysis were developed in the course of Emotech: Emotion and Technology Project undertaken at the University of British Columbia. This wide range of methods and materials served to generate an understanding of users’ learning experiences in HCI context in great detail. The essential feature of these approaches to data collection and analysis were the use of the video-stimulated recall in interviews conducted immediately after participants’ interaction with the five modules of Lucy to obtain participants’ reconstructions of their learning experiences and their emotional responses to Lucy. The screen display combining the images recorded by

the participant's camcorder was used for the purpose of the participant video-stimulated reconstructive interviews.

The main sources of data in this study were from participant reconstructive interviews, videotaped interactions, and observation field notes. Other sources of data included participant written work.

#### **3.6.1.1 Digitizing of videos data**

Data were collected at the time that participants interacted with each module of Lucy by using camcorders. Videotapes were digitized after participants' interaction with each module of Lucy for the purpose of participant reconstructive interviews and transcription. All video files, once digitized, were copied onto a DVD, together with all transcripts, scanned participants' written work and pre-research interview data.

#### **3.6.1.2 Participant reconstructive interview data**

Every participant was interviewed for about 60 minutes as soon as possible subsequent to his/her interaction with the five modules of Lucy. A video camera was employed during the interview to collect the interview data. The interview data were transcribed by using Transana and was copied to the DVD, together with other data as well as coded by HyperResearch.

#### **3.6.1.3 Observation field notes data**

Field notes in this study were very simple, for my roles in the lab were twofold: researcher and video-grapher. I recorded the time and types of all changes in the instructional activity, facial expressions, gestures and communication. The field notes table is shown below:



Table 3.3: Field Notes

Change Time	Instructional Activity	Facial Expressions	Gestures	Communication
00:00				
09:10				

#### 3.6.1.4 Participant written work during the interaction

All written work produced by every participant during the interaction with Lucy was photocopied immediately after each interaction and returned to the participant. Participants were required to bring their written work for the interview. An analysis of participant written work was undertaken after the interview, so the written work of all participants were asked to label with their names, the date, and the module of Lucy.

#### 3.6.1.5 Integrated data set

In summary, data triangulation and methodological triangulation were used in this study in order to reduce the likelihood of misinterpretation and thus strengthen the validity of the study (Patton, 2002). Video recording of participants' interaction with chatbot Lucy, the documentation of participants' work and participant reconstructive interviews provided a sophisticated account of thoughts of participants rather than researcher's interpretation.

In all, data collection in this study generated the following data set:

- ❖ Initial data set related to each interaction
  - Video tape from each participant camera
  - Digitized screen video files

- Video tape of each participant reconstructive interview
  - Observation field notes
  - Photocopies of participant written work produced during the interaction
- ❖ Additional general data set
- Pre-research interview

### **3.6.2 Data Analysis**

Five stages of data analysis were employed in this study: reviewing the whole data sets; identifying major constituent parts of data sources; focusing on human-computer interactions of each individual; coding the data set; comparative analysis of users' learning experiences (Erickson, 1991).

I started the analysis by reviewing the data sources: video tapes from each participant camcorder; digitized screen video files; video tapes of each participant reconstructive interview; observation field notes and photocopies of participant written work produced during the interaction. These data were digitized and transcribed. While transcribing, I identified major constituent parts of these data: images, video files, written files and transcriptions and began coding data set by using HyperResearch. Based on the theoretical framework described in chapter two, I adopted a two-level approach – emotional level and affective state level to code the data. At the level of coding, I paid attention to each individual interaction with Lucy. Last, but not least, I made a comparison of learning experiences of each participant to find similarities and differences in using chatbot technology in learning a foreign language.

### **3.7 Conclusion**

In this chapter, I discussed the methods and issues involved in the research design of this study. Informed by a theoretical framework of Evans' ontology of emotions, Scheffler's cognitive emotion theory as well as Krathwohl's affective domain of learning, this study adopted mixed methods and ethnographic methodology in order to capture details of users' learning experiences in the HCI context. The data collection of this study focused on users' learning experiences with the chatbot Lucy and their emotional responses to her. By triangulating a range of techniques in data collection, this study generated an integrated data set that included videotapes of actual interaction, observation field notes, videotapes of participant reconstructive interviews and participants' written work. These multiple data sources allowed me to understand users' (learners') learning experiences from various perspectives, which provided the bases for my analyses of interactions between users and Lucy. The analyses of data are reported in the next chapter.

## Chapter Four: Data Analysis and Findings

The last chapter described the data collected for this study and focused on users' (learners') experiences with chatbot Lucy. In this chapter, I address the questions:

- 4 How did users (learners) feel when using the chatbot Lucy to learn English?
- 5 How did users' (learners') feelings influence their cognitive processes?
- 6 Which objectives could users (learners) achieve in the affective domain of Lucy's learning model?
- 7 How could new language educational interfaces be designed to provide better support for users (learners) performance?

In answering the above questions, I first describe the approaches adopted to analyze the data, and then I report multi-layered analyses of users' experiences shown in Table 4.1. Users' learning experiences with the chatbot Lucy are analyzed through the lens of emotional learning and cognitive emotions and users' learning goals are described. New language educational interfaces are recommended.

### **4.1 A Narrative Approach to Analysis**

I employed narration to analyze the data. Narration most often involves the use of the past tense. While writing about users' experiences in using Lucy, I found writing in the present tense helped an easier recall of the tiny details that served to bring the story to life. I believed that when we narrated in the present we relived the moment. When we recalled something, we were distant from the emotions that were associated with an event. Hence, I used the present tense to write the narrative analysis of users' experiences with the chatbot Lucy.

Table 4.1: Five Stages of Analysis

Theme	Description
1. The Sensory Theme	By focusing on sensory engagement with Lucy, I showed how users' senses shaped their learning with the chatbot; in discussing sensory-learning processes, I analyzed how users' senses of the chatbot made their learning intrinsically meaningful for them; by measuring users' sensory accomplishment of levels in the affective domain, I discussed what level in the affective domain of learning users could achieve.
2. The Emotional Theme	By focusing on users' emotional responses to Lucy, I showed how Lucy's emotion generation influenced users' emotional reaction; through mapping users' emotions on the affective axes, I analyzed how users' emotions made their learning intrinsically meaningful for them; by measuring users' emotional accomplishment of levels in the affective domain, I discussed at what emotional level Lucy could facilitate users' affective learning.
3. The Cultural Theme	By focusing on users' cultural background, I showed how culture influenced users' experiences when using chatbot technology.
4. The Relational Theme	By focusing on the relations between users and chatbot, I showed how aesthetic and usable quality of the chatbot reflect the way in which users and chatbot relate to each other and how this relationship shapes their learning processes.
5. The Linguistic Theme	By focusing on the content and grammar embedded in chatbot, I showed how linguistic feature influenced users' feelings in using a chatbot tutor.

I began the analysis by examining the video recording of interactions of five chatbot modules and participant constructive interviews. All the interactions and interviews were transcribed using *Transana*<sup>TM6</sup>. The transcripts were created with an emphasis on detail: for example, gestures, facial expressions and other non-verbal behaviors were noted in brackets. While transcribing, I began coding the interactions using *HyperResearch*<sup>TM7</sup>.

This study used an open coding technique (Strauss, 1987). This process involved carefully reading the transcripts to determine the categories and themes. The constant comparative method allowed me to integrate the data collected (five interaction transcripts) and the theory employed (emotional learning and cognitive emotion). Some data related to the settings were incomplete due to the limitation of possible camera angles.

My analysis of the sensory theme, emotional theme, cultural theme, relational theme and linguistic theme showed the diverse dimensions of users' experiences with chatbot Lucy. The distinctions made among these five themes were not to privilege one over the other, nor to isolate the importance of a particular theme in the service of CALL interface design. Rather, this analysis aimed to identify the contributions of these five themes as a whole to the accomplishment of designing a user-centered multimodal adaptive language-learning interface.

---

<sup>6</sup> Transana<sup>TM</sup> software is for professional researchers who want to analyze digital video or audio data. See <http://www.transana.org/>

<sup>7</sup> HyperResearch<sup>TM</sup> is a qualitative data analysis software package enabling researchers to code and retrieve, build theories, and conduct analyses of their data. See <http://www.researchware.com/hr/tour.html>

Two levels of analysis were employed: level 1 – affective states and level 2 – emotional responses. Affective states and emotional responses were shaded within the process of a user's (learner's) interaction with their environment.

This study adopted Dewey's accounts of learning, meaning that an "individual learns in consequence of his/her direct activities" (Dewey, 1924, p. 199). "Learning is active. It involves reaching out of the mind. It involves organic assimilation starting from within" (Dewey, 1902, p. 9). In this study, I did not test the participants (users, learners) in order to know whether or what they learned, and how well they learned, but instead allowed the use of self-evaluation methods.

Hence, the users' experiences coding system consisted of three formatted tables (see Table 4.2, Table 4.3, Table 4.4). Table 4.2 provides the coding system of users' (learners') experiences and their related affective states. The left column of the table includes five components of users' experiences when using the chatbot. The top row lists users' affective states, which were categorized based on Kort's axes of emotional sets (see Chapter Two). Table 4.3 provides the coding system of users' affective states and their emotional responses to chatbot Lucy. The left column of the table includes users' affective states (results from Table 4.2). The top row listed users' emotional responses, which were categorized based on Evens' emotion spectrum: basic emotions, higher cognitive emotions and cultural-specific emotions.

Table 4.2: Users' Experiences - Affective States

Users' (Learners') Experiences	Affective States				
	Anxiety-Confidence	Boredom-Fascination	Frustration - Euphoria	Dispirited -Encouraged	Terror-Enchantment
Sensory					
Emotional					
Cultural					
Relational					
Linguistic					

Table 4.3: Users' Affective States - Emotional Responses

Affective States	Basic Emotions						Higher Cognitive Emotions						Cultural –specific Emotion
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	
Anxiety-Confidence													
Boredom-Fascination													
Frustration-Euphoria													
Dispirited-Encouraged													
Terror-Enchantment													



Table 4.4: Users' Experiences - Emotional Responses

Users' (Learners') Experiences	Basic Emotions						Higher Cognitive Emotions						Cultural -Specific Emotion MianZi
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	
Sensory													
Emotional													
Cultural													
Relational													
Linguistic													

Table 4.4 provides the coding system of users' experiences and their emotional responses to chatbot Lucy. The left column of the table indicates users' experiences. The top row listed users' emotional responses characterized by Evans' emotion spectrum: basic emotions, higher cognitive emotions and cultural-specific emotions.

## **4.2 *Participants Selection in Narration***

This research was conducted in two labs at a university. Fifteen participants were interviewed accordingly. However, not all these fifteen interviews were used for investigating the interface design. I used narrative analysis to articulate the story in order to capture the nuances in each interaction. Hence, I set up three criteria for choosing what data would be selected:

1. Interview data reported should inform the theoretical framework.
2. Interview data should be comprehensive enough to allow a range of analytical frameworks.
3. Interview data should be specific to the research design.

Based on these three criteria, five users' distinct experiences with Lucy were analyzed and five themes emerged. It did not mean that only one theme was relevant to one use. On the contrary, each user's experiences – sense, emotion, culture, relation, linguistics, were overlapping during his/her learning processes. However, for the purpose of accentuating each theme, I analyzed each user's peculiar experience.

## **4.3 *Lee's Sensory Experiences***

I will begin the narrative analysis of Lee's sensory experiences by briefly describing Lee's personal background and the research setting. I develop a description of

Lee's sensory experiences of interactions with Lucy's five learning modules in the lab. I analyze Lee's sensory-emotional processes, which focused on how Lee's sense of chatbot Lucy shaped his learning with chatbot Lucy. Finally I discuss the objectives embedded in Lucy that shaped Lee's experiences (see Table 4.10). Although this single account was not representative of all users' interactions with the chatbot, it drew on sensory aspects of moments of interaction that I found salient with respect to users' experiences with Lucy's interface.

### **4.3.1 General Background of Lee**

Lee is a male Chinese international graduate student in a technology program. He has been learning English since he was at grade five. He has been in Canada for a year, however, his English did not improve very much. He studies English in a language school in Vancouver, but due to the high tuition and his busy schedule in the program, he will not continue his English learning in the language school.

He was referred to me when I initiated the chatbot language interface research. He had experiences in using the practice and drill language learning software before and wanted to try Lucy's interface. He hoped that Lucy could provide him a different learning experience from the previous language-learning software he used. His goal in participating in this research was to improve his communication ability, especially to generate sentences in different contexts. He was very comfortable in using computers due to his bachelor's degree in electronic engineering. He spent many hours every day interacting with his computer – working, emailing, playing games and surfing the Internet.

### **4.3.2 The Research Setting**

Lee's interaction with Lucy started in October 2007 in one of the two labs used in this story (see Chapter Three). Five modules of Lucy were installed on two PC computers in the research lab and only Lee and I were in this lab each trial of the study. Lee had a quiet place to interact with chatbot Lucy without interruption and disturbance. In what follows, I paid attention to Lee's sensory experiences when using Lucy in this lab.

### **4.3.3 Lee's Sensory Experiences with Lucy**

Lee sits in front of the computer with Lucy installed, starting the program (Hotel Module), clicking on "Enter Lucy's World", watching her for a while. Without knowing how to make Lucy talk to him, he begins to try every icon on the screen. Lucy's eyes are blinking, her head is moving left and right and her hands knocking on the desk, which makes Lee puzzled. After five minutes, he tried all the icons. Now he starts his talking with Lucy using the snowball microphone. Lucy greets him, and he answers back by clicking on the "Talk" icon – "I'm going to need a room, please," which is shown at the bottom of the screen. Two seconds later, an 85% score appears besides the sentence he speaks. A smile is on his face. He exclaims: "Aha...this is exciting!"

He shuts down the music and continues his practice. He explains that the music is annoying during learning. He can't see any relationship between the music and the learning context and the repeated rhythm makes him bored. He prefers a quiet place to practice English.

Suddenly, a 100% score is shown on the screen, which makes him stands up.

"This is great! Aha... I get 100% score! I always doubt about my pronunciation."  
he smiles and says.

He completes one step of the conversation without any difficulty and starts over.

Lee starts another conversation with Lucy using the same module. However, he encounters some difficulties - Lucy can't understand what he says. He repeats the sentence several times, and Lucy keeps saying: "I didn't understand what you said. Try saying it again." He frowns again. He says the same sentence again, but Lucy can't understand him.

"What happened? What is wrong with my sentence? Is there a particular word that I made a mistake or is there something wrong with my intonation?" he repeats these questions twice.

He tries to change the intonation, and now Lucy understands him. "Aha..." he nods, "intonation, right, it is a question, so I need to use a rising tone." He smiles again. Joy can be detected from his facial expressions.

He keeps practicing the conversation, but he is stuck again. He tries repeating the sentence again and again, but Lucy doesn't understand him.

"What is wrong? I don't think my intonation is wrong. Maybe the pronunciation? But which words? If Lucy could let me know what was wrong with my conversation, I would benefit a lot from the correctness."

Suddenly, Lucy says: "I didn't hear you. Check your microphone and try again." Lee smiles and nods: "Good, I am happy that you tell me what is the problem." He checks his microphone and finds it disconnected from the computer for some reason. He reconnects the microphone to the computer and starts the conversation again.

He continues to practice. He nods continuously during his learning processes. In the following weeks, he comes to the lab regularly and starts the practice more and more

enthusiastically. He keeps practicing for three months and completes all five modules at the end of December 2007.

#### **4.3.4 The Sensory Theme**

It is through sense organs and five senses– touch, smell, hearing, taste and sight that human beings participate directly in the world around them. We sense our situation through our vision, hearing, somatic sensation (touch), taste and olfaction (smell). In HCI, vision, hearing and touch are weighted with high information processing potential and are seen as high bandwidth senses for communication between people and computing systems. These three senses – vision, hearing and touch also dominate educational software systems. In comparison, the senses of taste and smell are somewhat low data processing potential senses and few HCI systems engage them. As our senses are triggered, our feelings are aroused.

##### **4.3.4.1 Sensory – Learning Processes**

Lee's sense of sight and hearing with Lucy firstly influenced his affective states (see Table 4.5), which shaped emotional responses (see Table 4.6) to Lucy during learning. His positive emotions (see Table 4.7) triggered during learning determined his willingness to use Lucy's interface. Kort's model of emotions in learning was employed to identify how Lee's affective states influenced his learning processes to point out the weakness of existing language educational interfaces and investigate what chatbot technology could provide to language learners. Although modeling emotional axes is far from a settled question in Psychology and Cognitive Science, there is extensive evidence that we all experience at least some emotions while engaged in the learning processes.

My goal here is to construct a coherent model of the interplay of Lee's affective states and his learning.

Lee's emotional axis in learning (see Table 4.8) was divided into two halves – a positive valence half and a negative valence half. He experienced affective states such as frustration, boredom, discomfort and dissatisfaction on the one hand, and comfort, interest, satisfaction and excitement on the other. That was to say, he was obliged to mix the bitter and the sweet.

In this part of analysis, I firstly discuss Lee's emotions triggered during the interaction, and then I shed light on the interplay of his affective states and his learning.

Table 4.5: Sense of Sight and Hearing - Affective States

Senses	Affective States				
	Anxiety-Confidence	Boredom-Fascination	Frustration - Euphoria	Dispirited -Encouraged	Terror-Enchantment
Hearing	Discomfort	Boredom	Frustration	Dissatisfaction	
Sight	Comfort	Interest	Frustration	Satisfaction	Excitement

Table 4.6: Affective States - Emotional Responses

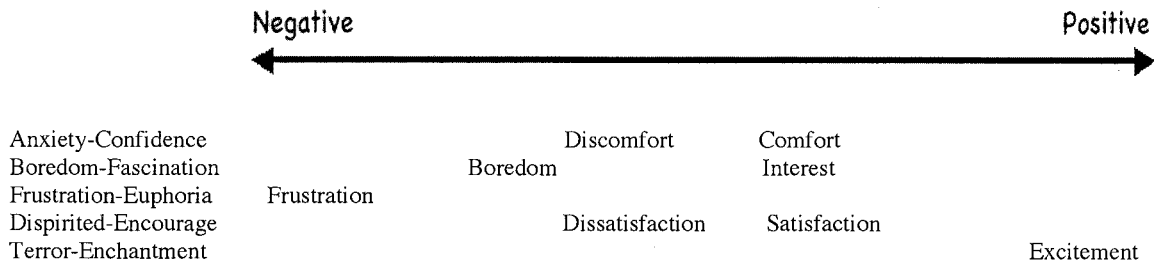
Affective States	Basic Emotions						Higher Cognitive Emotions						Cultural -specific Emotion Manzi
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	
Discomfort		Hearing											
Boredom		Hearing											
Frustration		Hearing Sight											
Dissatisfaction		Hearing											
Comfort					Sight								
Interest					Sight								
Satisfaction					Sight								
Excitement					Sight								

Table 4.7: Senses - Emotional Responses

													Cultural -Specific Emotion
Basic Emotions													Higher Cognitive Emotions
Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	Shame	MianZi
Sense	Hearing Sight			Sight									



Table 4.8: Axes of Lee's Emotional Sets in Learning



#### 4.3.4.1.1 Joy

At Lee's first sight of Lucy, he was very excited and happy to see the colorful background and animation as well as animated but human-like avatar Lucy. His feelings were completely different from the previous drill-practice learning software that he mentioned in the pre-research interview.

At my first sight of Lucy, I am excited to see Lucy and would like to have her help me learn English. It is more like a game. The colorful background situates myself in a particular learning context. Ah, I realize that she is different... You know, the powerful vision makes me excited.

His happiness and excitement about Lucy helped him continue to use the chatbot to learn in the following weeks. The power of his 'excitement' made Lucy an attractive option for learning. The colorful background and animation associated with green, blue, yellow and red influenced his feelings of comfort, interest, satisfaction and excitement, which contributed to his positive emotion – joy. The humanoid chatbot interface was welcomed by Lee.

I feel very comfortable when using the interface. The color and the animation are well – designed. I feel excited when a good score shown beside the sentence that I say.... I am satisfied with her graphical design.

#### 4.3.4.1.2 Disgust

The background sound interrupted Lee's learning experience. When he first interacted with the interface, Lee felt uncomfortable and bored about the music. He didn't like the music, as it disturbed Lucy's understanding of his speech and his understanding of Lucy's speech. Therefore he turned the music off. He felt more comfortable learning in a quiet environment. On the other hand, the music had only one rhythm and repeated again and again, which made him feel bored. The unrelated music became noise during his learning.

Lucy's idle behavior, which consisted of eye blinking, head movement and hand gesture when she waited for Lee's input made him puzzled, for Lucy's head and hand movement seemed random and Lee didn't know the reasons that Lucy emitted these gestures at a given point in time.

Not surprisingly, Lee was likely to discontinue interaction due to the noisy music, but the "turn off" function allowed him to turn off the music. This "turn on" and "turn off" function helped Lee feel satisfied with Lucy's interface design.

I like the way that they allow me to customize the background music. I know there may be some other people who would like to have the music, but it is annoying for me. I feel very comfortable without it. The "turn on" and "turn off" function is good.

Last but not the least, the immature speech recognition technology led to frustration and dissatisfaction in Lee.

A real person will react to my sentence quickly. Lucy needs to take some time to react to what I am saying and can't provide me detailed

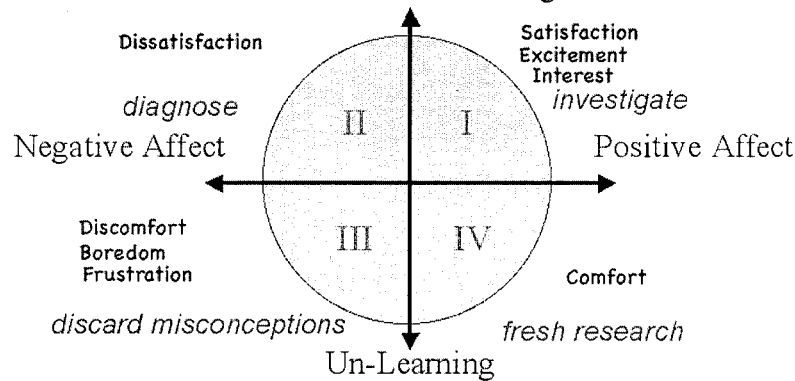
feedback of my pronunciation problems due to the limitation of speech recognition technology. Also, Lucy sometimes does not know what I am saying and I don't know what is wrong with me, which is not good for a language learner.

Affective states – discomfort, frustration, boredom and dissatisfaction shaped Lee's negative emotions - Disgust. The same rhythm of background music, Lucy's random body gestures and her embedded immature speech recognition technology frustrated Lee (see Table 4.7).

#### **4.3.4.1.3 Lee's Learning Processes**

Lee had a strong visual sense of Lucy's interface and of the interaction between himself and Lucy. The colorful interface made him feel comfortable; the background animation made him feel interested as if he was involved in a real conversation context; Lucy's gestures made him feel frustrated. He also had a strong hearing sense of Lucy and their communication. The background music made him uncomfortable and bored; the limited feedback and unrecognized speech made him feel frustrated and dissatisfied. Going beyond this, I would like to dig deeper into the interplay between his affective states and his learning cycle when using Lucy's interface. I employed Kort's comprehensive and scientific model to analyze the interplay between Lee's affective states and his learning (see Table 4.9).

Table 4.9: The Model of Interplay between Lee's Affective States and His Learning Cycle:  
Constructive Learning



Based on Kort's model of emotions in learning, Lee's affective states – satisfaction, interest, and excitement made him investigate what was embedded in Lucy and how Lucy was going to help him. He tried each button and function in order to determine how Lucy worked. At this stage, Lee had positive valence affective states (like excitement) as he investigated, explored and began to build his initial mental models for language learning. Eventually he had enough ideas in his head to be able to anticipate how Lucy was going to teach him and how he was going to learn. However, he realized something dissatisfied him, something he didn't predict. He needed to diagnose the discrepancy between his belief and what he heard. What was really happening was unrecognized speech and annoying background music as well as Lucy's random body gestures. He needed to make a decision about how he was going to deal with them. He tried to navigate Lucy with functional buttons, which could reduce his frustration. He shut down the background music and tried to ignore her gestures. He tried to be patient with Lucy's embedded speech recognition system. Eventually he discarded his beliefs that Lucy could be like a real human teacher. Due to his technical background, he understood Lucy's limitation.

I understand the limitation of technology, especially the limitation of speech recognition systems. My major is electronic engineering and I know it

is difficult for computers to accurately recognize users' voices and accents.

Although this made me feel frustrated, for I can't figure out what is wrong with my pronunciation, it is understandable.

He moved to the next learning stage – fresh research: he rediscovered how he was going to use Lucy in his learning and how Lucy was going to help him. He realized Lucy could be an English-speaking friend rather than a language teacher:

I like communicating with Lucy. She is more like a friend rather than a teacher, and she is so patient that I can practice my pronunciation with her as many times as I can. The library interface makes me recall the drill-practice software that I used before and the vocabularies, sentences and grammar that I learned in the class. I don't mean drill-practice is bad, and sometimes, it is helpful. Lucy's library is different in design, which allows me to compare my voice with those of the sample speakers.

After several interactions, Lee knew Lucy much better and he was interested in communicating with Lucy and felt satisfied for his learning. His positive sensory involvement made his learning intrinsically meaningful for him.

The beautiful graphic design of the interface and the vivid animation as the background situates myself in a context, such as at a hotel, in a restaurant or in a travel agency. Communicating with her is fun!

A holistic sense engagement could be seen in Lee's case, with a whole person participating in the interactions, his body and his mind, his senses and his intellect as well as his intelligence and his feelings. Lee's sensory engagement in Lucy and his motivation

to use Lucy overcame the frustration that the speech recognition system and random body gestures introduced.

Through examining Lee's case, background colors, animation, humanoid features and sound made Lucy an emotional entity. Lee's sense of seeing and hearing operated at the subconscious level and determined his affective states when interacting with Lucy during his learning processes (see Table 4.5). His affective states influenced his emotional responses to Lucy (Table 4.6). However, because of the understanding of Lucy's unavoidable limitation of speech recognition technology, Lee could tolerate the frustration and dissatisfaction, and thus his frustration and dissatisfaction evolved into disgust instead of anger during the learning process. The customized function – turning on and off function, enabled Lee to turn off the background music. As a result, he was willing to complete all five modules and constructed her as an English-speaking friend rather than a language teacher, for his language learning in the future.

#### **4.3.4.2 Sensory - Affective Domain of Learning**

The affective instructional domain examined here aimed to provide directions for designers to assess whether or not users were learning what was intended and built into the learning models.

Receiving is the first level in the affective domain (see Table 4.10), which deals with attention. Regarding Lee's sense in his learning processes, he received information through looking at Lucy's interface, reading what was shown on the interface and listening to what Lucy said to him and the background music.

Table 4.10: Sensory - Affective Domain:

	Receiving	Responding	Valuing	Organization	Characterization
The Sensory Theme	Looking at Lucy's interface	Displaying satisfaction & smiling	Preferring to interact with Lucy	N/A	N/A
	Listening to the background music and Lucy's saying	Turning off the background music & paying attention to Lucy's sayings	Declining to use the background music		

The responding level is active participation, which involves attention (receiving) and reaction. Lee's reaction to Lucy was due to his need of a quiet environment, so he turned off the music and paid attention to Lucy's words. He was satisfied with talking to Lucy because of her humanoid character but dissatisfied with Lucy's random body gestures (See Table 4.10).

When examining other data, I found all users were at the receiving and responding levels of the affective domain regarding their senses. All of them sensually participated at these lower levels.

It is much more difficult to assess students at the higher levels of the affective domain (valuing, organization and characterization). However, through the interview, I tried to assess users' higher levels of affective domain when using Lucy. Valuing is the value that is attached to objects, people or processes. During the interview, Lee mentioned that he preferred to interact with Lucy but declined to use the background music. For him, the toughest decision of continuing to use Lucy was how to conquer his frustration and boredom caused by speech recognition technology.

I can control my feelings of frustration and boredom toward the speech recognition system and understand the errors and limitation of that technology. I regard Lucy more as friend than as a teacher. A teacher should provide his/her student detailed feedbacks on problems occurred. But Lucy can't do that. However, she can become my good friend who can speak English.

Lee's evaluation of chatbot Lucy gave her less credit of being a language teacher but more credit of being a friend. This evaluation was concluded by his senses of Lucy: sight and hearing. The limitation of speech recognition technology played a decisive role in placing a valued Lucy.

#### **4.3.4.3 Decision Making – Does the Technology Helps?**

In Lee's case, emotions generated through his sense of chatbot Lucy influenced his evaluation of Lucy and his choice of whether or not to continue to use this technology in his learning. When Lee participated in the research, he was presented with an innovative technology toward which he might feel attracted.

Lee was reluctant to choose Lucy as his language tutor and prefers to have her as an English-speaking friend. He didn't think Lucy could teach him, but on the contrary, she could become a companion for language learners. The immature speech recognition technology was the biggest obstacle for Lucy to be a language tutor. For teaching the rules of a language, it is probably less important to have the humanoid features, such as eye blinking or body gestures. However, if the interface aimed to teach human-like and authentic communication, the humanoid features are important for providing communicative behavior that is "real".



The speech recognition technology is still in its developmental process. It is not a mature technology, for I realize that during my learning, it can't figure out my mistake. Sometimes, I make a mistake, but Lucy can still pass the conversation, and sometimes I don't make any mistake, but Lucy can't understand what I am saying. If it is a learning interface, it should let learners know clearly what the problem is and how to correct it. Also a learning interface needs to distinguish the right and wrong. A language learning interface is not a conversational agent that is for fun but it is for helping us correct problems.

The emotional responses of disgust toward Lucy's quality might lead Lee to stop using her. However the aesthetic graphic designs made Lee build a connection of friendship with Lucy, which made Lee continue to use Lucy as a friend who could talk to him in English.

I don't think Lucy can be my tutor, but she can become a friend who would like to communicate with me in English. You know, here in Canada, I don't have many local friends and due to my poor English, I feel very nervous when I talk to people who speak English. Lucy can make me feel relax when I talk to her in English. I don't need to care about much of grammar and pronunciation when talking to her. If she can't understand, I can repeat it. I believe this can also be called learning.

### **4.3.5 Summary**

In this section, I investigated Lee's emotional response to Lucy, in which I hoped to explore the weakness of Lucy's learning interface. I analyzed the sensory-learning processes, in which I focused on how Lucy helped Lee in his learning and Lucy's affective domain of learning. I concluded the discussion with Lee's decision making – whether or not he would like to use Lucy to learn English.

Lee's frustration with the speech recognition system embedded in Lucy and his joy toward Lucy's interface were inextricably linked with his learning and shaped how he valued the chatbot.

## **4.4 *Nu's Emotional Experiences***

I begin the narrative analysis of Nu's experiences by briefly describing the background of Nu and the setting. And then I shed light on the description of Nu's experiences with five modules of Lucy in the lab. I analyze Nu's emotional processes when interacting with Lucy and address Lucy's affective features. Finally I explore which level Nu achieved in Lucy's affective domain of learning (see Table 4.16) and how Nu valued Lucy to help him learn English. Although this single account was not representative of all interactions with chatbot Lucy, it drew on Nu's affective aspects of moments of interaction that I found salient with respect to his experiences with Lucy's interface.

### **4.4.1 General Background of Nu**

Nu is a male Thai international exchange graduate student studying science education. He also pursued an ESL class in the international house at the university, where his English level was defined as intermediate. He used computers in his language

learning before, and he was good at using technology for working or playing games. He has been in Canada for half a year and his English improved very fast.

When I met him, I mentioned the advantage of using the chatbot technology and chatbot language tutor Lucy for ESL learners. By showing him a short video of Lucy, he was enthusiastic to participate in my research. He hoped that Lucy could help him learn English expressions in contexts.

#### **4.4.2 The Research Setting**

This study was conducted in the same lab as Lee's beginning November 2007 (see 4.4.2) and the learning environment was the same. In the following, I present Nu's emotional processes when using Lucy and how he experienced affective features of Lucy.

#### **4.4.3 Nu's Emotional Experiences with Lucy**

Nu carries a notebook and a learning schedule for using chatbot Lucy. He plans his time schedule of using chatbot Lucy and how much time he is going to spend on each module. He writes his goal of using chatbot Lucy: remembering new phrases. He is very well organized in planning his learning.

Nu skips orientation sessions and enters into "Lucy's World" directly. Lucy, dressed as a receptionist, welcomes him: "What can I do for you?" Background music is playing. Nu clicks the 'Talk' button and speaks one of the three sentences shown on the text message: "I'm going to need a room, please." However, Lucy replies: " I couldn't understand what you said, try saying it again." Nu tries again, but Lucy still doesn't understand him. Nu changes another sentence in the message box, and Lucy doesn't understand him either. He tries six times before he turns to the researcher: "What happens? It does not work!"

He does not know what he is going to do and he looks at Lucy. He clicks her and suddenly she repeats herself, which scared Nu: “Oh, there is a repeat function! It scared me!”

He cannot figure out his problems of communicating with Lucy. He tries each of them and Lucy can’t understand all of them. He is so disappointed with Lucy. He says: “Poor me! What is wrong with my pronunciation? If Lucy can tell me which word has a problem, I may learn; but if not, I don’t know what I am going to do. I feel upset.”

He clicks the sample voice on the right side of each sentence and tries to imitate them. After a while, he tries to click the ‘Talk’ button again and Lucy answers him as well as a 98 percent score is shown besides what he says.

“Why 98 percent? I don’t know what is the difference between the sentences that I speak before and after listening to the sample. I am confused.”

He continues his talking with Lucy. Sometime the score is low and sometimes the score is high. He gets 100 percent twice but he shakes his head. It seems that he does not believe the score.

“I like the way they design a human-like language tutor who can communicate with me in English, but I don’t like the score system, for it is not accurate. I know my pronunciation, and I don’t think I can get a 100 percent score. Sometimes I realize some problems in my pronunciation, but Lucy can still give me a high score; and sometimes I don’t think I make any mistake about my pronunciation, but Lucy doesn’t understand what I am saying.”

Although he has some frustration and distrust feeling toward Lucy, he completes the conversation designed in each module of Lucy’s World.

In the following week, he pays attention to the library and tutorials session and spends more time on these two parts than other sessions. He talks to Lucy very quickly and if Lucy cannot understand him, he exits Lucy's world and then enters it again. He completes all the modules at the end of January 2008. He is the one who spends the longest time using Lucy among the participants.

#### **4.4.4 The Emotional Theme**

The emotional theme in this study responds to users' emotional responses to Lucy's ability to address<sup>8</sup> users' emotions. Educational technology can no longer be seen as only a powerful cognitive tool or medium, providing to the learner a non-judgmental and patient tutor. "The burden of adaptation has gradually been shifting from the human user to the computer" (Hudlicka, 2003, p. 2). Based on this, the design of Lucy's interface aims to decrease the users' frustration by implementing a humanoid character with facial expressions, gestures and body movement. However, Lucy's ability to generate her own emotion but inability to recognize and respond to users' emotions increased their frustration instead. Through examining the emotional theme, I aimed at discovering whether affective computing was a meaningful concept or a reasonable goal for educational interfaces to enhance users' learning.

##### **4.4.4.1 Emotion-Learning Processes**

Today there are a number of efforts to give computers certain emotional abilities. However, a long-term argument points out that "rather than trying to make computers (or computing) affective, we should try to make communication effectual. Rather than trying

---

<sup>8</sup> Computer's ability to address user's emotions refers to its ability to recognize user affect, adapt to the user's affective state, generate affective behavior by the machine, or model user's affective states.

to reproduce emotions, we should try to imitate those aspects of emotions that are known to enhance the effectiveness of communication” (Hollnagel, 2003, p. 70). The most controversial point is whether emotion enhances human-computer communication and therefore whether it should be incorporated into HCI (Hudlicka, 2003). Through examining Nu’s case, I investigated how Lucy’s emotion generation influenced Nu’s affective states (see Table 4.11), which shaped his emotional responses to Lucy (see Table 4.12). And then, I discussed Nu’s affective states when interacting with chatbot influenced his learning processes. Chatbot Lucy’s unsuccessful upper body movement, gestures and facial expressions made Nu’s learning unsuccessful. Nu’s emotional responses to Lucy’s emotion generation (see Table 4.13) determined his unwillingness to use chatbot technology in his language learning.

#### **4.4.4.1.1 Joy**

Lucy in these five modules had ability to express her emotions when she talked or when she waited for users. Her emotions were shown through her facial expressions, head movement, eye blinking and hand knocking. Lucy’s emotional expressions during her interactions with Nu made Nu feel frustrated like Lee. Nu’s hopefulness triggered his positive emotion – joy. Such emotional communication provided Nu emotional feedback on how well he performed, besides the score feedback. On the other hand, Lucy’s emotional feedback allowed Nu to easily understand whether his articulation could be understood or not. This was important for him and generated ‘Joy’ in using Lucy. It seemed to Nu that Lucy was a real person. She had her own emotions. When she understood him well, she smiled, which made Nu feel hopeful with his learning and when she didn’t understand Nu, she was frustrated, which at the beginning made Nu feel

curious about how to make Lucy understand him. But when he couldn't figure out how to reduce Lucy's frustration, he felt distressed.

#### **4.4.4.1.2 Distress/Disgust/Anger**

However, Lucy's inability to recognize and adapt to Nu's emotions made him distressed, and then this distress after a while became disgust and finally created anger toward Lucy, which decided his unwillingness to use Lucy as his language learning tool in the future. Lucy's five modules cannot sense, recognize and adapt Nu's emotions. Nu had to adapt Lucy instead of being adapted, and thus the emotional communication in the long run was interrupted by unrecognized emotions, which caused Nu to feel angry. Lucy's frustrated expressions made Nu understand that something was wrong with his pronunciation, but continuous frustration on her face and her gesture caused Nu feel distressed and he didn't know what he could do. After a while of distress, Nu felt that he didn't want to use Lucy anymore, for she couldn't provide any emotional understanding of Nu. Thus, Lucy's frustration at this point made Nu feel disgust instead of distress. If at this time, Lucy still couldn't figure out Nu's disgust (in fact, Lucy couldn't), and she still couldn't recognize what Nu was saying, Nu became agitated and finally angry.

Lucy's continuous frustration makes me nervous and anxiety. I want to correct myself in some way, and I try it again and again but I fail. Lucy doesn't provide me any hint on my mistake. What she does is only to say – "I didn't understand what you said. Try saying it again," however, I want to know what is wrong with me. Her frustration doesn't help me much but make me feel distressed. I don't like her.

Table 4.11: Lucy's Emotion Generation – Nu's Affective States

Lucy's Emotion Generation	Nu's Affective States					
	Anxiety-Confidence	Boredom-Fascination	Frustration - Euphoria	Dispirited -Encouraged	Terror-Enchantment	
<b>Frustration</b>	Anxiety	Curiosity	Frustration	Dissatisfaction		
<b>Happiness</b>	Hopefulness					

Table 4.11: Nu's Affective States – Nu's Emotional Responses

Nu's Affective States		Basic Emotions					Higher Cognitive Emotions					Cultural -Specific Emotion		
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	Shame	MianZi
Anxiety			Lucy's Frustration											
Frustration		Lucy's Frustration												
Dissatisfaction	Lucy's Frustration													
Hopefulness					Lucy's Happiness									
Curiosity					Lucy's Frustration									

Table 4.12: Lucy's Emotion Generation - Nu's Emotional Responses

	Basic Emotions										Higher Cognitive Emotions							Cultural -Specific Emotion
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	Shame					
Lucy's Emotion Generation	Lucy's Frustration	Lucy's Frustration	Lucy's Frustration		Lucy's Happiness									Míanzi				



In Nu's case, the consideration of Lucy's ability to recognize emotions was critical in order that Nu could successfully complete learning tasks, avoid errors and achieve optimal performance as well as maintain reasonable stress levels. An adaptive learning interface, which can adapt to individual affect and performance, is necessary. Lucy's failure to adapt to Nu's emotions made him feel angry finally in his learning processes. Hence, the affective factors need to be considered in this particular learning context and it is critical that designers should accurately assess the range of possible affective states users (learners) may, or should experience during interactions with the learning interface.

#### 4.4.4.1.3 Nu's Learning Processes

Nu's emotional axis in learning (see Table 4.14) was divided into two halves – a positive valence half and a negative valence half. He experienced affective states such as anxiety, frustration and dissatisfaction on the one hand, and hopefulness on the other. He mixed these affective states in his learning processes.

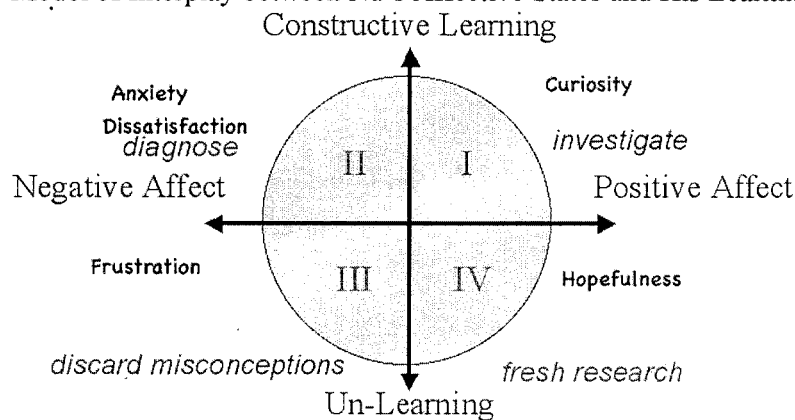
Table 4.14: Axes of Nu's Emotional Sets in Learning

	Negative	Positive
Anxiety-Confidence	Anxiety	Hopefulness
Boredom-Fascination		Curiosity
Frustration-Euphoria	Frustration	
Dispirited-Encourage		Dissatisfaction

Lucy's emotion triggered Nu's positive valence: Lucy's smile made him feel hopeful; However, Nu found an unconquered problem: Lucy's hand gesture aroused his anxiety with what he did; Lucy's head movement and eye blinking made him feel frustrated; and Lucy's facial expressions made Nu dissatisfied (Lucy couldn't understand what Nu said and showed a frustrated face to him, with her hands knocking on the desk

continuously). Nu felt anxious at the beginning and tried his best to clearly pronounce. But Lucy still didn't understand what he said, which made Nu feel dissatisfied with Lucy's design and he became frustrated, for he didn't know what he could do in order to be understood. Lucy's inability to sense and recognize Nu's frustration made him angry and he even wanted to quit from learning. Going beyond this, I mapped the interplay between Nu's affective states and his learning cycle when using Lucy (see Table 4.15).

Table 4.15: The Model of Interplay between Nu's Affective States and His Learning Cycle:



When communicating with Lucy, Lucy's inviting smile gave Nu incentive to continue his practice. He was curious about investigating Lucy's functionality and learning content. At this stage Nu directed his learning and accumulated the knowledge learned from Lucy. However, Nu found that Lucy randomly showed emotions as feedback to users but not in the sense that she recognized users' emotions. Lucy's failure to recognize Nu's emotions made him feel anxious and dissatisfied. He wanted to figure out what was wrong with his pronunciation. Unfortunately he couldn't get the answer. His anxiety became frustration. He felt frustrated in using Lucy, but he still tried his best to diagnose his problems until he found a solution. He felt hopeful again. After several times trials, he gave up and expressed his dissatisfaction with Lucy. When he felt anxious and dissatisfied, he still reported that he could learn, for he needed to find out a solution

to his pronunciation problem. His failure to figure out problems prompted him to discard his misconception. He continued to the next learning cycle, and after he tried all the five modules, he said that he wouldn't use the chatbot for his language learning.

#### 4.4.4.2 Emotion - Affective Domain of Learning

Lucy's emotion generation helped Nu achieve two lower levels in the affective instructional domain: receiving and responding (see Table 4.16). At the receiving level, Nu sensed Lucy's emotional expressions such as her happiness and frustration. Lucy's emotional expressions triggered positive and negative emotions when using her as a language Tutor. He also responded to Lucy's emotional expressions. When Lucy smiled, Nu felt hopeful for he realized he got a correct answer; when Lucy was frustrated, Nu's feeling changed from anxiety to dissatisfaction due to Lucy's inability to recognize his emotions. The current design of Lucy could not engage Nu in a higher level in the affective domain of learning, which might become a future goal for redesigning language learning interfaces.

Table 4.16: Emotional-Affective Domain

	Receiving	Responding	Valuing	Organization	Characterization
The Emotional Theme	Sensing Lucy's emotional expressions –Happiness & Frustration	Responding to Lucy's emotional expressions – happiness & Frustration	N/A	N/A	N/A

#### 4.4.4.3 Decision Making - Does the Technology Helps?

In Nu's case, emotions were engendered through his emotional communication with Lucy, which influenced his choice of whether or not to continue to use this

technology in his language learning. When I presented Lucy to him, he felt fascinated with this technology. When he started his learning using Lucy, he was amazed by the gaming – like interface design. However, when he found the emotional communication to be one way instead of two ways, he lost his interests in using Lucy. Lucy’s inability to sense and respond to Nu’s emotions made him feel dissatisfied and frustrated during the communication, which finally became a decisive key for him to give up using this technology for his language learning. He suggested that clear emotional feedback was critical for language learning interfaces like Lucy’s to help users (learners) to learn.

#### **4.4.5 Summary**

The emotional theme explored Nu’s emotional responses to Lucy’s emotion generation, his learning processes and his affective domain of learning as well as his decision-making. Nu’s emotions during his learning with Lucy were greatly influenced by chatbot Lucy’s ability to carry out emotional communication. Vague emotional feedback presented by Lucy triggered Nu’s frustration and dissatisfaction, which became anger at the end and prompted him to stop using Lucy.

### **4.5 *Lulu’s Cultural Experiences***

I begin the narrative analysis of Lulu’s experiences by briefly describing her background and the setting. I then shed light on the description of Lulu’s experiences of interactions with five modules of chatbot Lucy in the second lab used in this story. I analyze her cultural-learning processes, which focus on how Lulu’s Chinese background influenced her learning processes. Following Lulu’s learning processes, I discuss Lulu’s affective domain of learning (see Table 4.22) and what was important for her to make a decision to continuously use Lucy as a language-learning alternative. Although this single

account was not representative of all interactions with Lucy, it drew on Lulu's cultural aspects of moments of interaction that I found salient with respect to her experiences.

#### **4.5.1 General Background of Lulu**

Lulu is a female Chinese student in an ESL program in an English Learning Institute. Her English level was defined as intermediate. She used computers in her language learning before, and she liked using technology in her English learning. She was looking for language learning software that can help her pronunciation when I met her in the English Learning Institute. After showing her a short video of Lucy, she volunteered to participate in my research enthusiastically. I conducted this study in one of her lab sessions.

#### **4.5.2 The Research Setting**

This study was conducted in an English Learning lab beginning October 2007. This lab is a place for Language learning at the university. Five modules of chatbot Lucy were installed on fifteen computers in the lab (each computer with five modules). Fifteen users and one instructor as well as one researcher were in this lab each time a trial was conducted. Fifteen users used Lucy at the same time, and thus users did not have a quiet place to interact with Lucy. In what follows, I present how Lulu's Chinese cultural background shaped her learning processes when using Lucy.

#### **4.5.3 Lulu's Cultural Experiences with Lucy**

Lulu sits in front of the computer with chatbot Lucy installed, starting the program, clicking on "Enter Lucy's World". She starts talking to Lucy. A sentence spoken by Lulu is shown at the bottom of the screen with a score beside it. Her score of each sentence is high, no less than 85 percent. She shuts down the background music and speaks loud in

order that her voice won't be disturbed by others. She completes one turn of the conversation and starts over.

I feel confident when I talk to computer, but feel nervous when I talk to a person. I think I have a difficulty in talking to real people, for I don't like the eye contact, you know. Lucy does not need the eye contact.

Lulu starts another conversation with Lucy using the same module. However, she encounters some difficulties: Lucy can't understand what she is saying. She repeats the sentence, and Lucy keeps saying: "I didn't understand what you said. Try saying it again." She says the same sentence again, but still Lucy can't understand her. She exits the program and enters into "Lucy's world" again. Lulu explains that she used similar software before: sometimes these kinds of software don't work well; but if users (learners) exit the program and enter into again, the program will work.

I like using language learning software to learn English for it will help me keep my Mianzi (not loose face). When I talk to a person, I feel embarrassed when making mistake.

Noise becomes a problem. Since she is in a classroom with fifteen students and each student uses Lucy at the same time, although she has an earphone on her head, she still has a problem hearing Lucy clearly. Due to the noisy environment, Lucy cannot recognize her voice accurately sometimes. However, she has an instructor in the classroom. Each time she finds her method (exit – enter) does not work, she raises her hand. The instructor comes. Lulu asks the instructor to listen to her pronunciation and intonation in order to make sure it is not her problem but Lucy's problem of

understanding. Although she cannot pass to the next step of conversation if Lucy cannot recognize her speech, she can start the conversation over.

She continues to practice. She nods continuously during all her conversations in the Hotel Module. In the following weeks, she comes to the class regularly and practices her English with Lucy. This study lasts four months before the term ends. All fifteen students try each of the five modules.

#### **4.5.4 The Cultural Theme**

The culture theme in this study refers to the role that users' cultural backgrounds plays during their learning with Lucy. Users' cultural background become more and more critical in designing universal user interfaces that can be used by all potential users with various cultural backgrounds. Lucy was designed to be used by users from different cultures. All Lucy's learning materials are translated into seven languages including Chinese, Japanese, Korean, Vietnamese, Russian, Spanish and Portuguese. But Lucy can only speak English, requiring users to do so as well. Users can choose their native language when they enter into Lucy's world, which allows users to understand the meaning of what Lucy says and what is shown on the screen. Based on this, the design of Lucy aims to help as many ESL learners as possible. However, users from different cultures feel differently when they use Lucy. Through examining the cultural theme, I aimed to investigate users' learning processes and the 'one-fits-all' design and learning outcomes of using such an interface.

##### **4.5.4.1 Cultural - Learning Processes**

Studies of cultural influences on emotional responses demonstrate both that culture has an effect on emotional responses and that culture influences emotional responses in

very specific ways. Ethnographic notions by people such as Karl Heider (1991), Catherine Lutz (1988), and Sulamith Potter (1988) suggested that cultural contexts and heritages influence individuals' emotional responses. Our cultural backgrounds influence how we feel. However, what is less clear is whether different cultural beliefs and values regarding emotions can be translated into tangible differences in the various aspects of emotions in HCI settings. If a user, like Lulu, of Chinese heritage, holds to the cultural ideal of emotional *Mianzi*, how can we conclude that her Chinese culture shaped her emotional responses to Lucy? Through her physiological responses and her self-reported emotional experiences, I suggest that it is important to think of the computer as a social actor as replacing human actors with computer actors should not change social rules, which regulate human's interaction with each other.

Lulu's Confucian belief that views emotional moderation and control as a means of maintaining harmonious interpersonal relations influenced her affective states (see Table 4.17), which shaped her emotional responses to Lucy (see Table 4.18). These cultural - emotional processes (see Table 4.19) influenced her learning and determined her willingness of using chatbot technology in her language learning.

#### **4.5.4.1.1 Mianzi (Face)**

Lulu's emotional responses to Lucy were complicated. Her several basic emotions such as joy and embarrassment were combining and working together to contribute to her final reported emotion – *Mianzi* (face).

There is an old saying in China – Men live for face as trees grow for bark. There are two basic aspects of face in Chinese society and social relations: *Lian* and *Mianzi* (Hu, 1944). *Lian* refers to an individual's moral character in the eyes of others. The second



aspect of face is *mianzi*, representing social perceptions of a person's prestige. *Mianzi* is earned through success in life, or attaining a high or respected social position. A loss of *lian* would result in a loss of trust within a social network, while a loss of *mianzi* would likely result in a loss of authority. According to Hu (1944), although Westerners have a concept of 'face' similar to *mianzi* (i.e. 'social prestige'), it does not have the strong moral implications of the concept of *lian*. Face is a driving force in social relations among the Chinese, and failures to show *lian* or *mianzi* bring dishonor, disgrace, and shame to oneself (Mascolo, Fischer, & Li, 2003).

Consistent with a Confucian cultural face framework, Lulu always kept silent in the classroom when learning English, which was not good for her learning. She was afraid of making mistakes and loosing her face in front of other students. When I introduced Lucy to her class as a potential way to learn English, she felt excited and told me that she was looking for this kind of learning software for some time. She believed that she could learn better by using software than in the classroom.

When using computers for my language learning, I am never afraid of making mistakes and loosing my face. I become more confident in speaking.

Although she was still in a classroom (the lab), she wore her earphone and started talking to Lucy loudly in English. She visited Lucy's library, tutorial sessions and Lucy's world. It seemed that she was quite active in her learning unlike how she described herself as a shy student who seldom participated in classroom practice and discussions.

I feel stressful when I speak English in front of my teacher and other classmates. However, if I am quite fluent with some sentences, I feel comfortable to speak it loudly. Therefore, Lucy helps me a lot when I am at

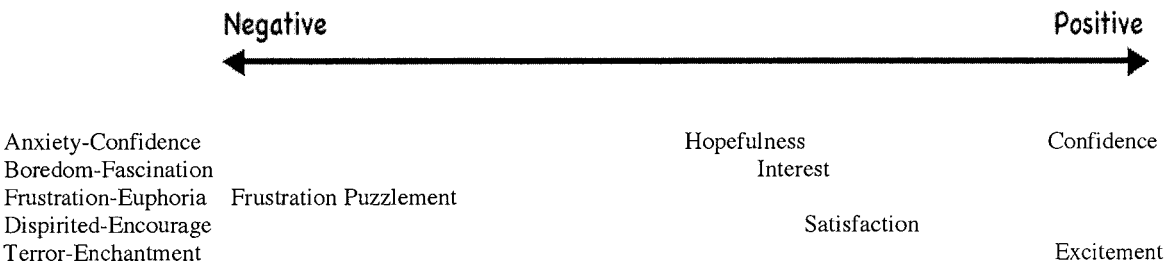
the beginning stage of my practice. She won't laugh at me... although I know teachers won't laugh at me too, I feel losing face when I stuck... when I stuck, I become nervous and can't speak anything...

Face is a cultural phenomenon that shaped Lulu's choice of using chatbot Lucy in her language learning. She prefers to have Lucy for her language learning because Lucy can save Lulu's face during learning. Being not afraid of making mistakes, Lulu was confident in speaking English loudly; she was interested in the content she learned and felt satisfied with her digital language tutor – Lucy; she is excited about the talking machine – Lucy, who was never tired of her mistakes.

4.5.4.1.2 Lulu's Learning Processes

Lulu's emotional axis in learning (see Table 4.20) was divided into two halves – a positive valence half and a negative valence half. She experienced affective states such as hopefulness, confidence, interest, satisfaction and excitement, which saved her face during her study with Lucy. She also experienced negative states such as frustration and puzzlement.

Table 4.17: Axes of Lulu's Emotional Sets in Learning





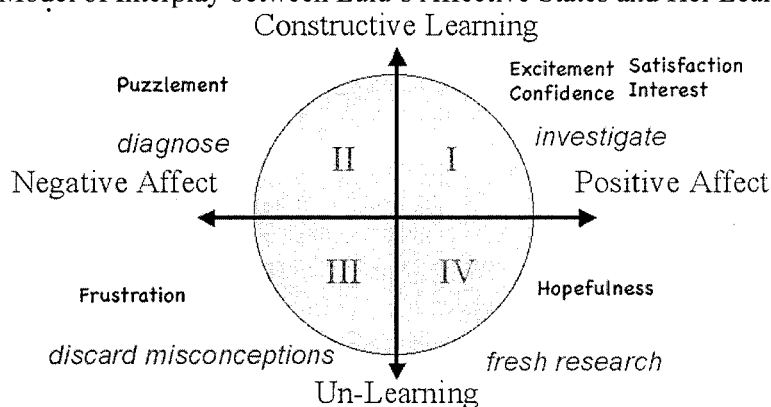
Lucy, as a language-learning tutor for Lulu, allows her to have the courage to make mistakes and can save Lulu's face during her learning processes. This is very important for Lulu to learn English. What she worried most about is her face during learning. She can't lose her face and she has to keep her *mianzi* when speaking English. Thus, having the ability of not making Lulu feel she is 'losing face', Lucy triggered Lulu's positive valence: Lucy's unrecognizable feedback triggered her interest in finding the mistakes she made; Lucy's score feedback made her excited, especially when she got 100 percent and she felt confident in speaking English. Although she also encountered the same technological problems as other users did, she tolerated these inconveniences and her frustration, for she thought Lucy could also tolerate her mistakes. Thus, she felt satisfied with Lucy. Going beyond this, I mapped the interplay between her affective states and her learning when using Lucy (see Table 4.21).

When interacting, Lucy's inability to recognize Lulu's speech sometimes triggered her interests in discovering what was wrong with her pronunciation and her speech. However, it was not all the time that she could get the answer, for sometimes due to the immature speech recognition technology, Lucy couldn't understand Lulu's sayings, which puzzled Lulu. But Lulu believed that her interests in talking to Lucy allowed her to research whether it was her mistakes or Lucy's inability to recognize her speech. She felt satisfied with Lucy's ability to tolerate her poor pronunciation and her hundreds of repetitions. Her interests in Lucy's speech recognition and satisfaction with Lucy's tolerance prompted her to investigate what Lucy could teach her. She started her learning. At this stage, Lulu constructed her learning and accumulated the knowledge learned from

Lucy. Every time when new features or functions embedded in Lucy appeared, Lulu felt interested and started her research. After the research, she continued her study.

During a three-month interaction, her physiological emotional responses and her self-report show that her positive affects were far more than negative affects involved in her learning processes. One of my initial conclusions was that she might experience more negative affects during her learning processes. However, she didn't want to mention them due to her face. When I examined her video clips of interactions, I found difficulty in seeing any emotions in her poker face except the excitement when she got 100 percent.

Table 4.21: The Model of Interplay between Lulu's Affective States and Her Learning Cycle:



#### 4.5.4.2 Culture - Affective Domain of Learning

Lucy's "one fits all" design helped Lulu achieve two levels in the affective instructional domain: receiving and responding (see Table 4.22). At the receiving level, Lulu received information coming from Lucy, such as her comments, her smile or frustration. Receiving Lucy's feedback triggered her interests and satisfaction in using Lucy as a language tutor. Lulu also responded to Lucy's conversations. All these saved Lulu's face during her English practice. She never felt she lost her face when she made a mistake. Lucy's response – "I couldn't understand what you said, try saying it again" allowed her to examine the problem by herself. Lulu's positive affects made her tolerate

Lucy's inability to recognize her speech. However, The design of Lucy couldn't provide Lulu a higher level of affective learning, which might become a future goal of redesigning language learning interface.

Table 4.22: Cultural - Affective Domain

	Receiving	Responding	Valuing	Organization	Characterization
The Cultural Theme	Receiving Lucy's information	Responding to Lucy's speech	N/A	N/A	N/A

#### 4.5.4.3 Decision Making - Does the Technology Helps?

In Lulu's case, emotions engendered through her interaction with Lucy influenced her choice of whether or not to continue to use this technology for her language learning. When I presented Lucy to her class, she felt excited in this technology. She was looking for this kind of technology for a long time. When she started learning by using chatbot Lucy, she was amazed by Lucy's ability to save her face. She never worried about her mianzi when practicing English pronunciation. She doesn't need to fear speaking English or making mistakes or losing her face in front of others. Her face became a decisive key for her to choose Lucy as her language tutor and tolerate Lucy's immature speech recognition system. She suggested was that Lucy could be better used at home instead of in the classroom due to the speech recognition technology, which works better in a quiet place than a noisy one.

#### **4.5.5 Summary**

The cultural theme explored Lulu's Confucius culture, which influenced her emotional responses to Lucy, and influenced her willingness in keeping Lucy as her private language tutor in the future. Lulu's cultural specific emotion – *mianzi* (face), which Lucy could save, greatly influenced her choice of using Lucy's learning interface. Face in this study became a decisive factor of whether Lucy would be chosen and kept for future use.

### **4.6 *Nim's Linguistic Experiences***

I begin the narrative analysis of Nim's experiences by briefly describing Nim's background and the setting. I then shed light on the description of Nim's experiences of interactions with five modules of chatbot Lucy in the lab. I analyze Nim's learning processes with Lucy's linguistic feature of each module, which focused on how Lucy's linguistic features influenced Nim's learning. Following Nim's learning processes, I attend to the goals that Nim achieved in Lucy's affective domain of learning (see Table 4.29) and what was important for Nim to make a decision to use this technology as a language learning alternative. Although this single account was not representative of all interactions with the chatbot tutor, it drew on the linguistic aspects of moments of interaction that I found salient with respect to Nim's experiences of using Lucy in her learning.

#### **4.6.1 General Background of Nim**

Nim is a female Thai international exchange graduate student studying science education. She also studied ESL in the international house at the university, where her

English level was defined as intermediate. She used computers in her language learning before. She has been in Canada for half a year and her English improved very fast.

When I met her, I introduced Lucy to her. She already knew Lucy due to Nu's introduction. She was enthusiastic to try Lucy. She hoped that Lucy could help her English pronunciation.

#### **4.6.2 The Research Setting**

This study was conducted in the same lab as Lee's and Nu's beginning November 2007 (see 4.4.2) and the learning environment was the same as theirs. In the following, I present Nim's learning processes with Lucy and her affective domain of learning.

#### **4.6.3 Nim's Linguistic Experiences with Lucy**

Nim starts her interaction with Lucy – Hotel Module by clicking “Library” in the main window. She says: “I choose to use the ‘library’ first because I notice that there is an explanation of ‘Library’ – ‘not sure how to make certain English sounds? Then this is for you’. This is my main purpose to use Lucy.”

She enters into Lucy's library. She chooses one of the buttons – voiced in a row to see the list of the sounds for this category. She selects /b/ sound from the list of sounds on the left to see the animations and video for this sound. She clicks on the play button of the video and imitates the pronunciation /b/. She clicks on the examples of /b/ sound: ball, taboo, and rub. She follows the sample pronunciation and practices the phonetics one by one. She says: “This is great! I can repeat as many times as I like. You know, some of the pronunciations are very difficult for Thai students to pronounce. However, by using the animations, I can see clearly how the sound can be pronounced.”



Satisfaction is shown on her face after she tries Lucy the first time. She doesn't have much time to try other parts in the same module. She asks me whether she can have more time to use this software for she is excited to have Lucy as her language tutor.

For the second time, she starts the tutorials session in the same module – Hotel Module. The tutorials session is for users (learners) to know how to improve their pronunciation on some difficult sounds.

Lucy, dressed as a hotel receptionist, welcomes her: “Welcome to our tutorials session.” Since Nim is not familiar with how to control the interface, she spends some time on being familiar with each button's function. She writes down her sequence of using chatbot: Clicking the next button and listening to how Lucy is explaining the pronunciation. Following the sequence of function buttons shown on the screen, she practices them one by one.

Lucy says: “Try saying the word [red].” Nim says: “[red].” She clicks the next button and Lucy says: “For a convincing /r/ sound, growl like a bear with your upper and lower teeth clamped and vibrating together, with your tongue kept well away from your teeth. Try saying [red] again.” Nim says: “[red].” Lucy says: “Can you feel the vibration?” Lucy continues to teach her /j/ sound. After Nim learns all the sounds in this session, she clicks ‘listen & learn’ button. There is no change on the screen. She clicks several times and no change. She then clicks “Hear the Difference” button and the screen changes. When she clicks the “Listen & Learn” button again, the screen changes back to the one that she used. Now she realizes that she completed the first part of this session. She sighs with relief. She moves to the second part in the tutorials session and plays with different sentences to train her ear to distinguish the difference between each sound. She

plays the sentence – That guy raps as well as anyone I know & That guy yaps as well as anyone I know, and listens to each sentence several times as well as reads each sentence. She smiles and nods. She completes the “Hear the Difference” part and says: “This part is excellent for me. I have a problem in figuring out the difference between close sounds. I like this part.”

She then continues to the “Listen & Choose” part. She clicks “Listen & Choose”, and plays the recording as well as chooses the sentence she hears. She clicks one of choices and a “yes” sound comes.

“This is exciting and it is very helpful.

She does very well on this part and gets all “yes” feedback. She smiles and says: “I like this software. It helps me to improve my pronunciation skills. I will buy one for myself.”

She clicks the “Speak & Score” button and plays the sample sentence, and records herself as what the instruction showed. She clicks the upper row of gray boxes to compare her speech to that of the native speaker. However, she is puzzled. It seems that she doesn’t understand something. She tries all ten sentences and then comments on this part: “I don’t think the sound wave as a feedback is useful for me. I don’t understand my sound wave. Should I have the same sound wave with the sample? I don’t think so. I am confused about the purpose of using the sound wave.”

She starts the practice session in the same module as the other day. When she enters into the practice session, Lucy says: “Here’s where you can use our speech evaluation software to compare your speech to that of a native English speaker.” Nim answers: “Really?” She clicks on the next button and Lucy says: “Each section is organized

according to language function, so you can focus your practice on the areas you feel are more useful to you.” She says: “Oh.” Lucy continues: “In each exercise, you’ll first listen to the native speaker, and then record your own voice.” Nim enters in. The interface is the same as what she uses in the “Speak & Score” part in the tutorials session. “What is the difference? Oh, I see. There is a step-by-step function in this part. However, in the “Speak & Score” part in the tutorials session, I also can practice each part of the sentence step-by-step by controlling the play button.”

In the following week, she pays attention to the library and tutorials session and spends more time on these two parts than other sessions. She talks to Lucy very quickly and if Lucy cannot understand her, she exits Lucy’s world and then enters it again. She completes all the modules at the end of January 2008. She spends the longest time learning English by using Lucy among the participants.

#### **4.6.4 The Linguistic Theme**

The linguistic theme in this study refers to linguistic aspects of each learning module in Lucy. Chatbot Lucy has five learning modules – Travel English, Helping Visitors, Restaurant English, Hotel English and Small Talk with each module focusing on each topic. Each module includes four parts: Enter Lucy’s world, which is a conversation component to help users practice daily conversations; Library, which allows users to practice some certain English sounds; Tutorials, in which Lucy teaches users how to improve their pronunciation of some difficult sounds and Practice, where users can get some extra practice before they talk to Lucy in Lucy’s world (see Table 4.23). Through examining the linguistic theme, I investigated Nim’s emotional responses to each component and how her responses influenced her learning.

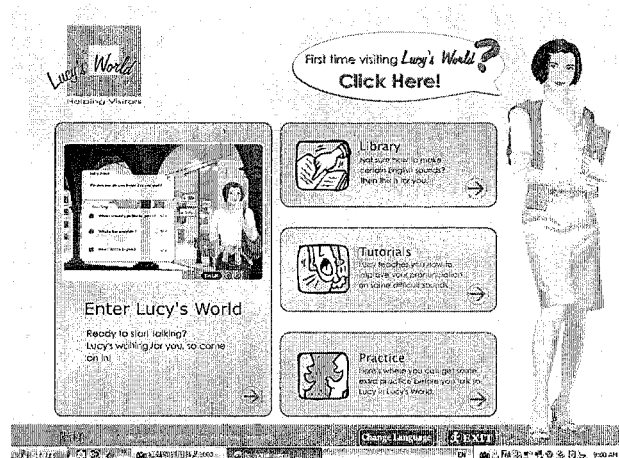


Figure 4: Four components of each Lucy's module

#### 4.6.4.1 Linguistic – Learning Processes

Chatbot Lucy is a language tutor, which has five linguistic modules with four components embedded in each module. The way that Lucy's each linguistic module was designed influenced the transmission of the linguistic content and shaped Nim's affective states (see Table 4.23) and her emotional responses to Lucy in her learning (see Table 4.24). These linguistic – learning processes determined her willingness of using Lucy in her language learning (see Table 4.25).

Table 4.23: Linguistic - Affective States

Affective States					
	Anxiety-Confidence	Boredom-Fascination	Frustration - Euphoria	Dispirited - Encouraged	Terror-Enchantment
Five Modules	Discomfort	Boredom		Dissatisfaction	
Four Components	Comfort	Interest	Puzzlement	Satisfaction	Excitement

Table 4.24: Affective States - Emotional Responses

Affective States	Basic Emotions						Higher Cognitive Emotions						Cultural -specific Emotion MilanZi
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	
Discomfort			Five Modules										
Boredom			Five Modules										
Dissatisfaction			Five Modules										
Comfort					Four Components								
Interest					Four Components								
Puzzlement			Four Components										
Satisfaction					Four Components								
Excitement					Four Components								

Table 4.25: Linguistic - Emotional Processes

Linguistic Aspects	Basic Emotions						Higher Cognitive Emotions						Cultural -Specific Emotion	
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	Shame	MianZi
			Five Module & Four		Four Components									

#### **4.6.4.1.1 Joy**

When Nim first interacted with Lucy, she was excited about the four components in each of Lucy's modules: Library, Tutorials, Practice and Enter Lucy's World. She spent most of her time on the library component, tutorial component and practice component. These three components allowed her to practice her pronunciations step-by-step, which satisfied her goal of participating in this study. She felt comfortable navigating each component and was interested in all the learning features in each component. She practiced so carefully and got excellent feedback, which made her feel satisfied with what she learned. The way that Lucy was designed in each component employed a comparison between recorded voice and a native speaker, which Nim found helpful and excited about.

Her positive affective states contributed to her positive emotional response to Lucy – Joy. She was happy to use Lucy for her pronunciation practice and was excited to know Lucy.

I like the three components – library, tutorials, and practice in Lucy. I feel very comfortable in using them. The comfort is very important for me during my learning, for when I feel nervous or anxious, I can't learn.... I am happy to have her.

#### **4.6.4.1.2 Distress**

However, Nim found that the design of five modules, with each module focusing on a different topic didn't fit for her well.

I don't like Lucy's World, for it doesn't look like a real life conversation. I understand the purpose of the design, which aims at setting up each conversation in a context. However, although in a restaurant context, we are

not necessarily required to talk about something only relating restaurant...

Lucy's world limits the conversation.

The separation of the conversation based on the context made Nim feel that this was unreal. The unreal conversation made her feel uncomfortable. Choosing one of the sentences shown on the screen made her feel like a talking robot, which triggered her boredom and negative affective state. She felt bored in using Lucy's world to practice conversation.

I don't think I can learn a lot from Lucy's world. All the sentences are fixed and I have to follow the way it was designed. It made me feel I am like a robot. In this way, I don't think it is appropriate for language learning.

Language learning should allow users construct sentences by themselves.

Nim was not only dissatisfied with Lucy's separation of the conversation (divided conversations into five modules) but was also dissatisfied with Lucy's inability to recognize sentences not shown on the screen. Her feelings – discomfort, boredom and dissatisfaction made her spend very little time in Lucy's world because these negative affective states triggered distress or negative emotion for her.

Also there are some overlapping designs among library component, tutorials and practice components, which made Nim feel puzzled. When she entered into a new component, she identified some already completed exercises and felt puzzled. She needed to go back and forth to figure out what happened and understood it was a design problem. The overlapping not only caused her puzzlement but also wasted her time in searching whether she did it or not completed these components. And thus, her distress also came

from her confusion toward the overlapping exercises among library, tutorials and practice components.

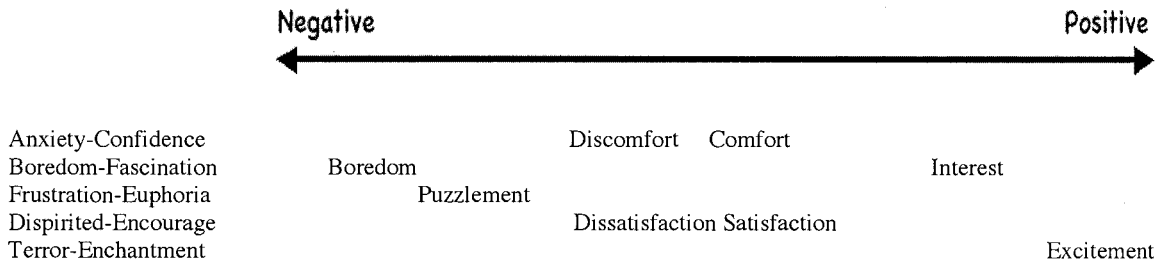
Through examining Nim's case, the design of each component in each module influenced Nim's emotions. Nim's pronunciation practice goal underwrote her love of the practice tool embedded in Lucy. However the overlapping exercise design puzzled her a little bit, but her love of these exercises overcame her confusion and she continued to use these exercises for her pronunciation practice. However since she believed that sentences should be generated and constructed by the speaker instead of imitation, the way that sentences were organized in Lucy's world could not be counted as language learning, she disliked the Lucy's world component. On the one hand, Lucy's world separated the real world conversation from conversations in real life, which made Nim feel it was unreal; on the other hand, Lucy's inability to recognize sentences not shown on the screen but generated by the speaker made the conversation mechanically, and Nim felt like a robot and bored.

#### **4.6.4.1.3 Nim's Learning Processes**

Nim's emotional axis in learning (see Table 4.26) was divided into two halves – a positive valence half and a negative valence half. She experienced negative affective states such as discomfort, boredom, puzzlement and dissatisfaction on the one hand, and positive affective states - comfort, interest, satisfaction and excitement on the other. She mixed these affective states in her learning processes.

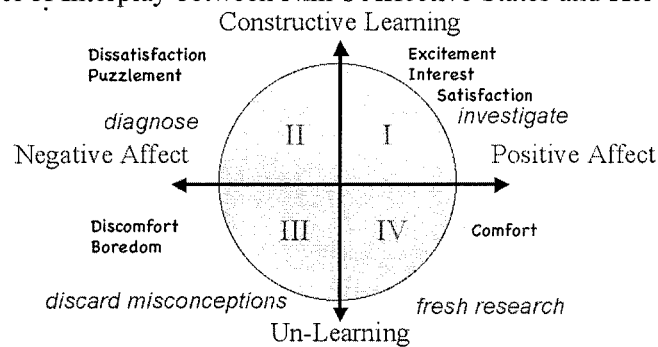


Table 4.26: Axes of Nim’s Emotional Sets in Learning



Lucy’s five separate modules triggered Nim’s negative valence; she felt she was like a robot, repeating something that was programmed; she was uncomfortable with conversation in a fake context and dissatisfied with the separate design modules; she felt bored after repeating fixed conversations again and again and puzzled about the overlapping exercises. However, the four components in each module helped her achieve her goal of practicing pronunciations and made her excited. She felt comfortable and interested in following Lucy’s explanations and satisfied with the interface design as well as her achievement. Going beyond this, I mapped the interplay between her affective states and her learning when using Lucy (see Table 4.27).

Table 4.27: The Model of Interplay between Nim’s Affective States and Her Learning Cycle:



Based on Kort’s theory of emotion and learning, Nim’s affective states – excitement, interest and satisfaction made her investigate the features in Lucy (four components in each module) and how Lucy was going to help her in her learning (five modules). She

started from the Library session, following Lucy's instructions step-by-step. At this stage, Nim had a positive valence affective state-excitement, interest and satisfaction as she investigated, explored and began to build her initial mental models for her language learning. Eventually she had enough ideas in her head to be able to anticipate how Lucy could teach her and how she was going to learn. However, she realized that something, such as overlapping exercises, made her confused and she felt dissatisfied with the design of Lucy. She tried to diagnose the problems and found it was a problem of design. After practicing, she felt bored and uncomfortable with the unreal fixed conversation in Lucy's world and quit from that component. Since her goal was to practice pronunciation, she continued the other three components, which made her interested and comfortable. She started fresh research on the other three components and her satisfaction with their design prompted her to investigate new knowledge embedded in Lucy.

#### **4.6.4.2 Linguistic - Affective Domain**

The affective instructional domain examined here aimed to provide directions for designers to assess whether or not users (learners) were learning what they intended and was built in the learning models.

Receiving is the first level in the affective domain (see Table 4.28), and deals with attention. Regarding Lucy's linguistic aspect of each learning module, Nim listened to what Lucy said to her. She received information through listening to what Lucy instructed and responded to Lucy by following what was shown on the screen. The design of Lucy couldn't provide Nim a higher level of affective learning, which might become a future goal of redesigning language learning interfaces.

Table 4.28: Linguistic - Affective Domain

	Receiving	Responding	Valuing	Organization	Characterization
The Linguistic Theme	Listening to Lucy's instructions	Following what was shown on the screen	N/A	N/A	N/A

#### 4.6.4.3 Decision Making – Does the Technology Helps?

In Nim's case, emotions engendered by interacting with Lucy's five modules influenced her choice of whether or not to continue to use this technology in her language learning. When she started her learning by using chatbot Lucy, she found that three components except 'Lucy's 'World' in each module helped her achieve her goal of practicing pronunciation, which made her excited and interested in using this software. Although during learning, she also felt puzzled by the overlapping exercise design and robot-like conversation that were disconnected with real life. However, she persuaded herself not to be bothered by some features that she didn't like in Lucy and continued to learn those she thought useful. This became a decisive key for her to choose Lucy as her language tutor and tolerate the overlapping design and separated conversations. She suggested was that designers should bring real life conversation to Lucy and be careful of the overlapping design in the learning modules. Instead of having five learning modules, Nim thought one learning module that could combine all five would be better. The conversation should be flexible instead of fixed.

### **4.6.5 Summary**

The linguistic theme investigated the content design of Lucy and Nim's learning processes with Lucy. Due to Nim's personal goal of pronunciation practice, she was quite satisfied with Lucy's linguistic learning module design. However, the issue of real life conversation design in this vignette was raised and influenced Nim's feelings toward using Lucy.

## **4.7 *Num's Relational Experiences***

I begin the narrative analysis of Num's experiences by briefly describing Num's background and the setting. I then shed light on the description of Num's experience of interactions with five modules of chatbot Lucy in the lab. I analyzed how Num's relations with chatbot Lucy influenced her learning processes. I also attend to her affective domain of learning (see Table 4.35) and what was important for her to make a decision to use the chatbot as a language-learning alternative. Although this single account was not representative of all interactions with the chatbot tutor, it drew on Num's relational aspects of moments of interaction that I found salient with respect to her experience when using chatbot technology.

### **4.7.1 General Background of Num**

Num is a female Thai international exchange graduate student studying science education. She is in an ESL class in the international house at the university, where her English level was defined as intermediate. She was good at using technology in learning. She has been in Canada for half a year and her English improved very fast.

I met her and introduced the chatbot Lucy to her. She already knew Lucy due to Nu's introduction. She was enthusiastic to try Lucy. She hoped that Lucy could help her English conversations.

#### **4.7.2 The Research Setting**

This study was conducted in the same lab as Lee's, Nu's and Nim's beginning November 2007 (see 4.4.2) and the learning environment was the same as theirs. In the following, I address the influence of Num's relations with Lucy on Num's emotional responses to Lucy and her affective states in relation to her learning.

#### **4.7.3 Num's Relational Experiences with Lucy**

When Num enters into Lucy's Travel Module, she clicks 'Enter Lucy's World'. Friendly Lucy greets her: "Is there anything I can help you with?"

"I'd like some information, please." Num says.

"Sure. What would you like to know?"

"I was thinking of visiting Europe." Num says.

"Sure. Are there any specific places you really want to visit?"

"What would you suggest?"

"When do you want to go?"

"I have time off in December."

"That's off-season, so the prices will be lower. But some places will be very cold."

"I don't mind the cold weather."

"In that case, perhaps a winter holiday in Sweden would be fun."

"That sounds delightful."

"Great. Let me get some more information for you."

“Start over.”

Num completes one step of the conversation. A smile is shown on her face. She starts over. She complete all the conversation in this module fast and instead of working on other components in this module, she asks me whether she can try other conversations in other modules. I agree with her request. She also tries the restaurant English and hotel English in her first time of interaction.

Oh, I like chatbot Lucy. I enjoy talking to her. She makes my learning interested and I feel happy.

In the following days, she continues to work on the conversation component of each module back and forth. Her desire is to be familiar with all the sentences shown on the screen. She believes that language learning at its beginning requires learners to recite and repeat. The recitation and repetition are meaningful for her. The new design of recitation and repetition makes her care for Lucy and is repaid by the enjoyment of recitation and repetition. For her, talking with the humanoid Lucy again and again doesn't make her bored. On the contrary, this interaction prompts her motivation to do the recitation and repetition.

She had a learning experience with language learning software before and she hates those tedious practice and drill. The short human-human conversation makes her excited.

I love the human – like chatbot Lucy. Although chatbot Lucy doesn't change its way to teaching – practice and drill, it has a humanoid feature, which makes me feel differently. It looks like I am talking to a real person. Lucy, to some extent, is more like a friend than a teacher.

With the humanoid feature designed for Lucy, Num feels personally committed:

Talking to Lucy, I don't feel I am talking to a machine.... I don't know... It is very strange for me. With Lucy's graphical design and animation, I feel I am talking to a person... not a computer... I feel a connection with Lucy.

#### **4.7.4 The Relational Theme**

The relational theme in this study responds to the influence of the relationship between Num and chatbot Lucy on Num's emotional responses to Lucy. This relational experience refers to "transactions between us and the objects and events that make up the world in which we act" (MaCarthy & Wright, 2004, p. 90). Through examining the relational theme, I explore how relationship configures into meaningful learning.

##### **4.7.4.1 Relational – Learning Processes**

Num's relational experiences with chatbot Lucy, which I described above and her experiences engaging in conversation of Lucy were readily characterized by her immersion in the transaction. She lost her sense of the separation of the language learning and herself. That is, when she was immersed in her learning experiences with Lucy, the elements of these experiences were so interdependent that she lost her sense of the separation of herself, Lucy and events of learning. Distinctions between these elements were highlighted when something went wrong with an experience or when she paused to reflect on the experience for some other reason. Therefore, she was a part of the learning system and her learning is a part of her. There was no gap to be bridged between her and chatbot Lucy's system.

Hence Num's relational experience with Lucy influenced her affective states (see Table 4.29), which shaped her emotional responses to Lucy (see Table 4.30). In turn, I

discuss how Num's affective states when interacting with the chatbot influenced her learning processes. Num's relational connection with Lucy determined her willingness to use chatbot technology in her language learning (see Table 4.31).

#### **4.7.4.1.1 Joy**

Num was happy with the chatbot conversation design. Her goal of using Lucy was to practice her conversational ability. She had an experience of using practice and drill language learning software, and she was amazed by the new way of thinking. She had a strong connection with Lucy, that is, she lost the sense of separation of the learning system and herself. She was hopeful about the humanoid Lucy and was interested in talking to her as if talking to a real person. She was satisfied with Lucy's interface design and the learning modules. She thought the context-based learning helped her become familiar with sentences that could be used in different settings, which made her feel excited. However, she also felt frustrated when Lucy couldn't recognize her voice. She stopped for a while and then changed options in order to make Lucy understand her and save time. She knew this was the problem of the speech recognition technology, however, due to her close relation with Lucy, she tolerated it. Her confusion came from Lucy's related feedback. It seemed that all the connections such as the microphone and computer worked very well, but Lucy asked her to check her microphone. When something went wrong (Lucy couldn't recognize Num's voice or error of the feedback) Num highlighted the problems within the design of Lucy's learning module rather than Lucy, therefore, she believed that humanoid Lucy would have a brilliant future. And thus, both her positive affective states and her negative affective states contributed to her emotion – joy.



Table 4.29: Relational - Affective States

Affective States					
	Anxiety-Confidence	Boredom-Fascination	Frustration - Euphoria	Dispirited - Encouraged	Terror-Enchantment
Relation	Hopefulness	Interest	Frustration/Puzzlement	Satisfaction	Excitement

Table 4.30: Affective States - Emotional Responses

Affective States	Basic Emotions						Higher Cognitive Emotions						Cultural –specific Emotion MianZi
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride	
Hopefulness					Relation								
Interest					Relation								
Frustration					Relation								
Puzzlement					Relation								
Satisfaction					Relation								
Excitement					Relation								

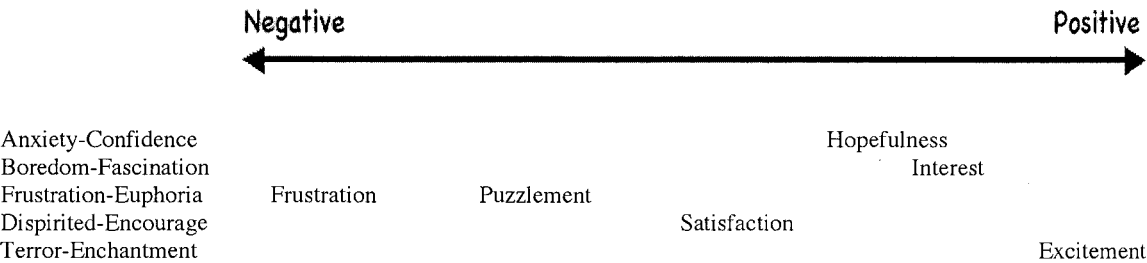
Table 4.31: Relational - Emotional Processes

Relation	Basic Emotions					Higher Cognitive Emotions							Cultural -Specific Emotion	
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	Love	Pride		Shame
					Relational Experience									

4.7.4.1.2 Num’s Learning Processes

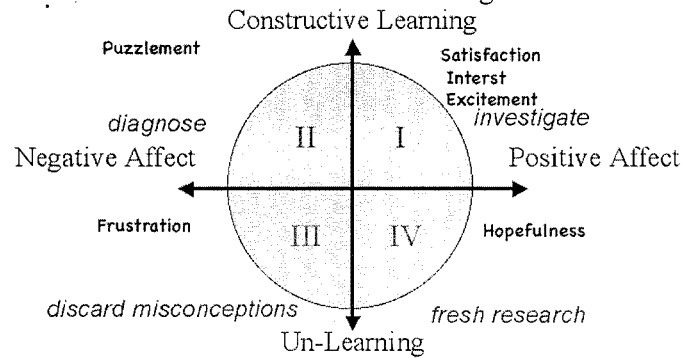
Num’s emotional axis in learning (see Table 4.32) was divided into two halves – a positive valence half and a negative valence half. She experienced negative affective states such as confusion and frustration on one hand, and positive affective states - hopefulness, interest, satisfaction and excitement on the other. She mixed these affective states in her learning processes.

Table 4.32: Axes of Num’s Emotional Sets in Learning



Lucy’s humanoid features triggered Num’s positive valence: she felt hopeful for the future of language learning software; she was interested in learning with Lucy and lost her sense of separation of herself from the learning; she was satisfied with Lucy’s conversation component design; and most important she was excited to have her as her language tutor. During the interaction, she also encountered some difficulties. However, she understood technology and could tolerate the inconvenience that technology brought to her. Her frustration and confusion didn’t stop her in using Lucy as her language tutor. Going beyond this, I mapped the interplay between her affective states and her learning when using chatbot Lucy (see Table 4.33).

Table 4.33: The Model of Num's Affective States in Learning



Based on Kort's theory of emotion and learning, Num's affective states – satisfaction, interest and excitement made her feel a connection with Lucy. She started from “Lucy’s world” and conversed with Lucy like a human conversation. She didn’t practice other components in each module except “Lucy’s world”, which made her excited. At this stage, Num had a positive valence affective state – satisfaction, interest and excitement as she investigated, explored and began to build her initial mental models for the conversation. Eventually she had enough ideas in her head to be able to communicate with Lucy. However, she realized a few things, such as immature speech recognition technology, made her puzzled and she felt frustrated with it. She tried to diagnose the problems and found it was a problem of technology. After a while of practicing, she felt a close relation with Lucy and didn’t feel that she was learning. She was just talking to someone in the restaurant and was ordering food. She found that humanoid Lucy was helpful in achieving her goal of practicing her conversational ability and remembering phrases for daily life. Num started a fresh research on each of “Lucy’s five modules in “Lucy’s World” and practiced them again and again. Her satisfaction with humanoid Lucy and her connection with her prompted her to investigate new conversations offered by Lucy.

#### 4.7.4.2 Relational – Affective Domain

The affective instructional domain issues examined here aimed to provide direction for designers to assess whether or not users were learning what they intended and built into the learning models.

Receiving is the first level in the affective domain (see Table 4.34), and deals with attention. Regarding Lucy's linguistic learning modules, Num started a conversation with Lucy. She received information through listening to Lucy and reading what was written on the screen shown inside Lucy's box. She responded to Lucy by talking to her. The design of Lucy couldn't provide Num with a higher level of affective learning, which might become a future goal of redesigning language learning interfaces.

Table 4.34: Relational - Affective Domain

	Receiving	Responding	Valuing	Organization	Characterization
The Linguistic Theme	Listening to Lucy's instructions	Following what was shown on the screen	N/A	N/A	N/A

#### 4.7.4.3 Decision Making – Does the Technology Help?

In Num's case, emotions that were engendered from interacting with Lucy's world influenced her choice of whether or not to continue to use this technology in her language learning. She became immersed in her learning and lost her sense of the separation of herself Lucy and then learning system. She did not feel she was learning in as much as she felt. However, she was communicating with someone who was in a context such as a hotel or restaurant. Although during learning, she also felt puzzled and frustrated by the

immature speech recognition technology, she understood it. Therefore, the immature technology did not influence her learning very much. She continued to learn what she thought was useful for her. This became a decisive key for her to choose Lucy as her language tutor and tolerate the puzzlement and frustration that occurred during learning. She suggested that the advance of speech recognition technology would improve the future of Lucy.

#### **4.7.5 Summary**

The relational theme explored how relations between Num and Lucy influenced Num's using chatbot Lucy and her learning outcomes. Her joyful feelings from talking to humanoid Lucy made her connected to Lucy in spite of the immature speech recognition technology. She believed that humanoid chatbot Lucy had a bright future in language education.

### **4.8 Conclusion**

This chapter documented the multi-layered analyses of five users' interactions with chatbot Lucy. These analyses delineated five themes that users experienced when using Lucy. Applying a lens of emotions and learning, this study offered insights into the nature of users' experiences with technology as it occurred in both the classroom lab settings and the individual lab learning settings.

First, I showed that the main weaknesses of Lucy's interface through investigating five users' experiences with Lucy. The speech recognition system was one of the most important problems that influenced users' choice of Lucy as a language tutor. Immature speech recognition technology made users feel bored and frustrated in using Lucy. Lucy's idling behavior, which consists of her eye blinking and her small head movements

while she waits for users' input as well as her random hand movement became a second weakness of Lucy's interface. Users had difficulty in figuring out the meaning of her visible gestures. The unadaptive Lucy's interface, which lacked the ability to recognize users' emotions, hindered users' learning.

Second, technology, especially speech-driven multimodal adaptive technology could help language learners learn a foreign language. The potential opportunity of the speech – driven multimodal adaptive interface lies in its interactive communication ability, which was important for language learning.

Finally, through examining the interplay between users' emotions and their learning processes, new language learning interfaces could take advantage of multimodality and adaptivity. In the next chapter, I discuss the significance of these findings and draw some implications to inform further research into interface design that could help students learn a foreign language.

## **Chapter Five: Conclusions, Recommendations And**

### **Summary**

The purpose of this study was to investigate user-centered speech-driven multimodal adaptive language learning interface design through the lens of users' experiences. This study focused on five users' experiences with chatbot Lucy. Data gathered for this study were designed to answer three research questions stated in section 1.4:

**Question 1:** How did users feel when using Lucy to learn English?

**Question 2:** How did users' emotional responses to Lucy's interface influence their cognitive processes?

**Question 3:** At what level in the affective domain of learning did Lucy facilitate users' learning?

This final chapter discusses the findings of this study in relation to the above research questions and the literature reviewed in Chapter Two. The discussion of implications was constructed under two headings: multimodality and adaptivity. Reflections on some limitations of this study were articulated, followed by some recommendations for future research. The final section provides a summary of this chapter.

## **5.1 *Conclusions for Research Questions***

### **5.1.1 How did users feel when using Lucy to learn English?**

#### **5.1.1.1 Users' Affective States**

Users experienced a wide range of affective states when using Lucy to learn English (see Table 5.1). Frustration and confusion in this study became the most experienced negative affective state of users and were major reasons why some of users couldn't reach their goals by using Lucy. Due to the frustration and puzzlement that users experienced, they were hesitant to use Lucy's applications and avoided using her for future language learning. The frustration and confusion, when investigated deeper, were found to be mainly caused either by the speech recognition system embedded or Lucy's idle behaviors. Some users spent nearly one-third of their learning time trying to figure out the speech recognition system and Lucy's random gestures.

Excitement was the most experienced positive affective state that users had when using Lucy. They felt excited with the graphic design, humanoid communication, and game-like buttons etc. Their excitement prompted their willingness to use Lucy and to continue their learning with Lucy.

Besides these two affective states, users experienced other feelings in their learning processes such as dissatisfaction or satisfaction, boredom or interest, anxiety or hopefulness, curiosity etc. These affective states worked together and shaped users' (learners') basic emotions like joy or disgust towards Lucy's interface (see Table 5.2).



Table 5.1: Users' Affective States

Users		Affective States					
		Anxiety-Confidence	Boredom-Fascination	Frustration - Euphoria	Dispirited -Encouraged	Terror-Enchantment	
		Discomfort	Boredom		Dissatisfaction		
Lee	Comfort		Interest	Puzzlement	Satisfaction	Excitement	
Nu	Anxiety / Hopefulness		Curiosity	Frustration	Dissatisfaction		
Lulu	Hopefulness Confidence	/	Interest	Frustration / Puzzlement	Satisfaction	Excitement	
Nim	Discomfort / Comfort		Boredom / Interest	Puzzlement	Dissatisfaction Satisfaction	Excitement	
Num	Hopefulness		Interest	Frustration / Puzzlement	Satisfaction	Excitement	

Table 5.2: Users' Emotional Responses

Users	Basic Emotions										Cultural -specific Emotion
	Anger	Disgust	Distress	Fear	Joy	Surprise	Embarrassment	Envy	Guilt	Jealousy	
Lee		Sensory			Sensory						
Nu	Emotional	Emotional	Emotional		Emotional						
Lulu											Cultural
Nim			Linguistic		Linguistic						
Num					Relational						

#### **5.1.1.2 Users' Emotional Responses to Lucy**

Users' emotional responses to Lucy were influenced by affective states triggered by Lucy. Joy was one of the emotions that occurred in the learning processes. Four users among five reported that they were happy and felt joyful in using Lucy, which was decisive in determining whether or not they were going to use Lucy in future learning. Disgust was another emotion reported by users in this study. Users' frustration, puzzlement, dissatisfaction, boredom and discomfort when using Lucy caused their disgust towards Lucy, which hindered their willingness to use Lucy for their future study. However, four users found that Lucy, to some extent could provide help as a learning companion rather than a language tutor. Therefore, they accepted her for their language learning. During the investigation, one cultural specific emotion – *Mianzi* was reported by one of the users. This cultural specific emotion was one of the merits that Lucy had and human teachers didn't have. Lucy helped users saved their faces when they tried new pronunciations, composed new sentences and expressed their ideas. Due to this contribution, Lulu preferred to have Lucy teach her English although she experienced many difficulties using Lucy. Anger and distress were reported by Nu, which were caused by long-time disgust. Nu's anger and distress finally made him give up using Lucy for his language learning.

#### **5.1.2 How did users' emotional responses to Lucy's learning interface influence their cognitive processes?**

Focusing on Kort's interplay between emotions and technology and cognitive emotions, this study shed its light on how users' emotions contributed to their learning processes. As indicated in Chapter Four, due to different affective states users had in their

learning processes, their learning was changed accordingly. When they felt excited they would continue to research more and at the same time, they were learning. However if users felt uncomfortable or frustrated in using some part of the application, they couldn't learn anything through communicating with Lucy.

### **5.1.3 At what level in the affective domain of learning did Lucy facilitate users' learning?**

Through examining five users' learning experiences, I concluded that Lucy's interface couldn't help users achieve high levels of affective learning. Lucy's objective in five modules was mainly to ask users to receive or respond to the information. Only relational and sensory experience helped users achieve a higher level – valuing.

In summary, the findings outlined above made it clear that when interacting with Lucy, users experienced sensory, emotional, cultural, linguistic, and relational engagement, which influenced users' learning processes and their choices of using her as a facilitator.

## ***5.2 Practical Implications***

### **5.2.1 Designing Multimodal Communicative Language Learning Experiences**

The findings reported in this study had important implications for multimodal communicative and integrative CALL applications. Language use was the key for most language teaching applications. If the interface was to be conversational and behave like a human, it must be able to reason and interpret multimodal information so that it can provide users can consistent and authentic feedback. As mentioned (see Chapter Four),

Lucy's learning interface employed a strategy of randomly emitting gestures or eye blinking instead of interpreting multimodal information. As a result, confusion and frustration occurred. Future interfaces should be endowed with more advanced conversational capabilities for reasoning behind non-verbal behavior. Authentic language use involving human-like communication will eventually meet the standard of human communication.

Secondly, a second language implies a second culture, which affected how a user interacted with the interface. So, the multi-cultural aspect was clearly very important, because understanding the culture was an important key to understanding the language use. Traditional second language teaching in school made language learners struggle with grammar rules and glossaries for a number of years without the learner being able to successfully communicate with native speakers in the culture of the target language. However, a useful communicative CALL system with the aim of teaching a second language could have potential to implement such features, although contemporary multimodal systems do not.

Last but not the least, one important form of communication is continuous feedback. Continuous feedback is difficult to be mimicked – much less produced in a random fashion – but can actually be generated by a knowledge inference system. This requires very fast and partial interpretation of user input on the fly. Since we don't know enough about how to quantify and model continuous feedback and today's modality recognition engines cannot handle such fast integration and interpretation, fast feedback is still a theoretical technology issue.

### **5.2.2 Designing a User – Adaptive Interface**

Lucy's interface offers users realistic opportunities for individual tutoring. Users can tailor Lucy for their own pace of learning: enter an answer to every question, repeat a lesson not fully comprehended or skip lessons that merely rehearse previously acquired skills. However, less experienced or less able users may not be able to make an accurate assessment of their own shortcomings or to devise for themselves a coherent and comprehensive study plan. For these users, Lucy can't adjust learning paces and emotions during learning, which cause some users discontinue interactions.

Future user-adaptive language learning interfaces should be able to change automatically in response to experiences or user performance. That new interfaces will change to suit the skills, emotions and knowledge of an individual user. Such a user-adaptive interface will no longer be constructed for a stereotypical average user, but for users' varying abilities or different skills by providing information and exercises at levels to match their current needs.

### **5.2.3 Methodological Issues**

For this study, I developed an approach to describe in detail users' experiences by employing narrative methods. There were two reasons that I used a narrative approach to analyze the data: 1) the process of narrative analysis explores details of users' experiences when interacting with Lucy. These nuances were what made one user's experience different from another. 2) Narrative analysis could help researchers and readers attain an understanding of users' experiences – the more the details recalled while narrating an incident, the more vivid was the description, offering an understanding and appreciation of the experiences.

Such a narrative approach investigating users' experiences made specific requirements of the analytical skills of the analyst as well as the techniques of data collection and transcribing. First, it required the analyst to carefully examine the details of users' practices in order to discover the factors that influenced their learning processes. It not only involved the examination of the human-computer interaction, but also visible and embodied acts such as users' facial expressions and think aloud utterances. Second, such analysis required the collection of data that maintained as much information as possible about the settings, and visual aspects of users' interaction, as well as their non-verbal communications. This meant making crucial analytical decisions about where to locate the camera in order to record the relevant data source. Third, this analytical approach posed special challenges for transcribing the data. In particular, it required the transcriber to make pragmatic decisions about what should and should not be included in the transcripts, and to be able to present the data in the format that maintained its rich details.

### **5.3 *Recommendations for Future Research***

By applying a theory of emotions in learning and cognitive emotions, this study generated some important findings for designing future CALL applications. In particular, it pointed to the importance of users' experiences – sensory, emotional, cultural, linguistic and relational, in the process of learning with technology. Given the model of factors influencing the design of chatbot language interfaces and factors influencing cognitive and emotional connections with the chatbot language tutor, a further investigation into the role of each factor would be extremely helpful: (1) Hold sensory, emotional, cultural and relational factors constant while the linguistic factor is

manipulated for understanding the implications of the linguistic content and database of the interfaces. (2) One chatbot could be designed as a language expert with extraordinary language skills and another chatbot could be designed as a language expert with feedback; a comparison could be studies between these two chatbot tutors and we could determine the importance of these variables in the creation of chatbot interfaces and user confidence. (3) Corresponding technical skills of users, which affect each factor addressed in the study, need to be further investigated.

## **5.4 Summary**

Overall, this study opened up possibilities for connecting emotions with cognition in users' experiences with technology in language learning. Approaches developed in this study could be useful in researching other users' experiences with technology. Insights gained from this study could also inspire more users' experiences research, which could further enrich language educational interface design.

## Bibliography:

- Abelson, R. P. (1963). Computer simulation of hot cognition. In S. S. Tomkins & S. Messick (Eds.), *Computer simulation of personality* (pp. 277-298). New York: Wiley.
- Abras, C., Maloney, D., & Preece, J. (2004). User-centered design. In W. Bainbridge (Ed.), *The encyclopedia of human-computer interaction* (pp. 656-659). Thousand Oaks: Sage Publications.
- Arnold, M. B. (1960). *Emotion and personality*. New York: Columbia University Press.
- Batson, C. D., Shaw, L. L., & Oleson, K. C. (1992). Differentiating affect, emotion and mood: toward functionally based conceptual definitions. In M. S. Clark (Ed.), *Review of personality and social psychology* (Vol. 13, pp. 294-326). Newbury Park, CA: Sage Publications.
- Bedford, E. (1957). Emotions. *Proceedings of the Aristotelian Society*, 57, 281-304.
- Bentley, R., Hughes, J. A., Randall, D., Rodden, T., Sawyer, P., Shapiro, D., et al. (1992). *Ethnographically-informed systems design for air traffic control*. Paper presented at the Conference on Computer Supported Cooperative Work, Toronto.
- Blomberg, J., Giacomi, J., Mosher, A., & Sewenton, P. (1993). Ethnographic field methods and their relation to design. In D. Dchuler & A. Namioka (Eds.), *Participatory design: principles and practices* (pp. 123-155). New Jersey: Erlbaum.
- Bogdan, R., & Biklen, S. (1992). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn & Bacon.
- Brentano, F. (1874). *Psychology from an empirical standpoint*. London: Routledge & Kegan Paul.
- Broad, C. D. (1954). Emotion and sentiment. *The Journal of Aesthetics and Art Criticism*, 13, 203-214.
- Campos, J. J. (1994). The new functionalism in emotion. *SRCD Newsletter*.
- Cannon, W. (1929). *Bodily changes in pain, hunger, fear and rage* (2nd ed.). New York: Appleton.
- Cohen, J. B., & Areni, C. S. (1991). Affect and consumer behavior. In Robertson & H. H. Kassarijan (Eds.), *Handbook of consumer behavior* (pp. 188-240). Englewood Cliffs, NJ: Prentice-Hall.



- Conati, C., & Zhou, X. (2002). *Modeling students' emotions from cognitive appraisal in education games*. Paper presented at the 6 International Conference on Intelligent Tutoring System, Biarritz, France.
- Cosmides, L., & Tooby, J. (2000). Evolutionary psychology and the emotions. In M. Lewis & J. Haviland-Jones (Eds.), *Handbook of emotions* (pp. 91-115). New York: Guilford Press.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.
- Darwin, C. (1872). *The expressions of emotions in man and animals*. New York: Philosophical Library.
- Deign, J. (1994). Cognitivism in the theory of emotion. *Ethics*, 104, 824-854.
- Derbaix, C., & Pham, M. T. (1998). For the development of measure of emotion in marketing: Summary of prerequisites. In M. Lambkin, G. Foxall, F. A. Raaij & B. Heilbrunn (Eds.), *European perspectives on consumer behavior* (pp. 140-155). London: Prentice Hall Europe.
- Dewey, J. (1902). *The child and the curriculum*. Chicago: University of Chicago Press.
- Dewey, J. (1924). *Democracy and education*. New York: Macmillan Company.
- Ekman, P. (1972). *Emotions in the human face*. New York: Pergamon Press.
- Ekman, P. (1973). *Darwin and facial expression*. New York: Academic Press.
- Ekman, P., & Friesen, W. (1969). The repertoire of nonverbal behaviour. *Semiotica*, 1(1), 86-89.
- Elliott, C., Rickel, C., & Lester, J. (1999). Lifelike pedagogical agents and affective computing: An exploratory synthesis. *Artificial intelligence today: Lecture notes in computer science 1600*(VIII), 195-211.
- Emde, R. N. (1980). Levels of meaning for infant emotions: A biosocial view. In W. A. Collins (Ed.), *Development of cognition, affect, and social relations: Minnesota Symposium on child psychology* (Vol. 13, pp. 1-37). Hillsdale: Lawrence Erlbaum Associates Inc.
- Erickson, F. (1991). Ethnographic microanalysis of interaction In M. D., W. LeCompte, L. Millroy & P. Judith (Eds.), *The handbook of qualitative research in education* (pp. 201-226). San Diego, California: Academic Press.

- Ericsson, K. S., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215-251.
- Evans, D. (2001). *Emotion*. New York: Oxford University Press Inc.
- Fetterman, D. M. (1998). *Ethnography: Step by step (2nd)*. Thousand Oaks, London, New Delhi: Sage Publications.
- Frank, R. (1988). *Passions within reason: The strategic role of emotions*. New York: Norton.
- Frijda, N. (1986). *The emotions*. New York: Cambridge University Press.
- Hammersley, M. (2006). Ethnography: Problems and prospects. *Ethnography and Education*, 1(1), 3-14.
- Harre, R. (1986). *The social construction of emotions*. Oxford: Basil Blackwell.
- Healy, J., & Picard, R. W. (2000). *SmartCar: Detecting driver stress*. Paper presented at the 15th International Conference on Pattern Recognition.
- Hochschild, A. R. (1983). *The managed heart: Commercialization of human feeling*. Berkeley, California: University of California Press.
- Hollan, J., Hutchins, E., & Kirsh, D. (1999). Distributed cognition: A new foundation for human-computer interaction research. *TOCHI Special Issue on Human-Computer Interaction in the New Millennium*, 7(2), 174-196.
- Hollnagel, E. (2003). Commentary: Is affective computing an oxymoron? *International Journal of Human-Computer Studies*, 59(1-2), 65-70.
- Horvitz, E. (1999a). *Attention-sensitive alerting*. Paper presented at the UAI '99, Conference on Uncertainty and Artificial Intelligence.
- Horvitz, E. (1999b). *Principles of mixed initiative interaction*. Paper presented at the CHI '99, ACM SIGCHI Conference on Human Factors in Computing Systems.
- Hudlicka, E. (2003). To feel or not to feel: The role of affect in human-computer interaction. *International Journal of Human-Computer Studies*, 59(1-2), p. 1-32.
- Hume, D. (1739). *A treatise of human nature*. Buffalo: Prometheus Books.
- Izard, C. E. (1971). *The face of emotion*. New York: Appeton Century Crofts.
- Izard, C. E. (1977). *Human emotions*. New York: Plenum.

- Jain, A. K., Duin, R., & Mao, J. (2000). Statistical pattern recognition: A review. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(1), 4-37.
- James, W. (1884). What is an emotion? *Mind*, 9, 188-205.
- James, W. (1950). *The principles of psychology* (Vol. 2). New York: Dover.
- Jorgensen, D. L. (1989). *Participant observation: A methodology for human studies*. Newbury Park, CA: Sage.
- Kaapor, A., Mota, S., & Picard, R. W. (2001). *Toward a learning companion that recognizes affect*. Paper presented at the AAAI Fall Symposium: Emotional and Intelligent II, the tangled knot of social cognition.
- Kaufman, L., & Weed, B. (1998). Too much of a good thing? Identifying and resolving bloat in the user interface: CHI 98 workshop. *SIGCHI Bulletin*, 30(4), 46-47.
- Kenny, A. (1963). *Action, emotion and will*. London: Routledge and Kegan Paul.
- Kerly, A., Hall, P., & Bull, S. (2007). Bring chatbots into education: Towards natural language negotiation of open learner models. *Knowledge-Based Systems*, 20, 177-185.
- Kort, B., Reilly, R., & Picard, R. W. (2001a). *An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion*. Paper presented at the IEEE Conference on Advanced Learning Technologies, Madison.
- Kort, B., Reilly, R., & Picard, R. W. (2001b). *External representation of learning process and domain knowledge: affective state as a determinate of its structure and function*. Paper presented at the Artificial Intelligence in Education Workshops, San Antonio, TX.
- Kuutti, K. (2001). *Hunting for the lost user: From sources of errors to active actors and beyond*. Paper presented at the Cultural Usability Seminar.
- Larson, J. A., Oviatt, S. L., & Ferro, D. (1999). *Designing the user interface for pen and speech applications*. Paper presented at the CHI'99 Workshop, Conference on Human Factors in Computing Systems.
- Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist*, 39, 124-129.
- Lazarus, R. S. (1991). *Emotion and adaption*. New York: Oxford University Press.
- Lazarus, R. S. (1999). *Stress and emotion: A new synthesis*. New York: Springer Publishing.

- Lazarus, R. S. (2001). Relational meaning and discrete emotions. In R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 37-67). New York: Oxford University Press.
- LeCompte, M. (1993). A framework for hearing silence: What does telling stories mean when we are supposed to be doing science? In W. Tierney & D. McLaughlin (Eds.), *Naming silenced lives: Personal narratives and the process of educational change* (pp. 9-27). New York, NY: Routledge.
- Lepper, M. R., & Chabay, R. W. (1988). Socializing the intelligent tutor: Bringing empathy to computer tutors. In H. Mandl & A. Lesgold (Eds.), *Learning issues for intelligent tutoring systems* (pp. 242-257).
- Leventhal, H., & Tomarken, A. J. (1986). Emotion: Today's problems. *Annual Review of Psychology*, 37, 565-610.
- Lincoln, Y., & Guba, E. (1985). *Naturalistic inquiry*. Beverly Hills, CA: Sage Publications.
- Lindgaard, G. (2004). Adventurers versus nit-pickers on affective computing. *Interacting with Computer*, 16(4), 723-728.
- Lyytinen, K., & Yoo, Y. (2002). Issues and challenges in ubiquitous computing. *Communications of the ACM*, 45(12), 62-65.
- McCarthy, J., & Wright, P. (2004). *Technology as experience*. Cambridge, Massachusetts, London, England: The MIT Press.
- Mascolo, M. F., Fischer, K. W., & Li, J. (2003). Dynamic development of component systems of emotions: Pride, shame, and guilt in China and the United States. In R. J. Davidson, K. Scherer & H. H. Goldsmith (Eds.), *Handbook of affective science* (pp. 375-408). Oxford: Oxford University Press.
- Mascolo, M. F., & Griffin, S. (1998). *What develops in emotional development*. New York: Plenum.
- McGrenere, J., Baecker, R. M., & Booth, K. S. (2002). An evaluation of a multiple interface design solution for bloated software. *CHI 2002*, 4(1), 163-170.
- Mikula, G., Scherer, K. R., & Athenstaedt, U. (1998). The role of injustice in the elicitation of differential emotional reactions. *Personality and Social Psychology Bulletin*, 24, 769-783.

- Millen, D. R. (2000). *Rapid ethnography: Time deepening strategies for HCI field research*. Paper presented at the Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, New York.
- Minsky, M. (1988). *Society of mind*. New York, NY: Simon & Schuster, Inc.
- Murray, I., & Arnott, J. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature of human vocal emotion. *Journal of the Acoustic Society of America*, 93(2), 1097-1108.
- Neu, J. (1977). *Emotion, thought, and therapy: A study of Hume and Spinoza and the relationship of philosophical theories of the emotions to psychological theories of therapy*. London: Routledge and Kegan Paul.
- Nisbett, R. E., & Wilson, T. d. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84(3), 231-259.
- Nussbau, M. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge: Cambridge University Press.
- Nussbaum, M. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge; New York: Cambridge University Press.
- Oakley, J. (1992). *Morality and the emotions*. London: Routledge and Kegan Paul.
- Oviatt, S. L. (1999). The myths of multimodal interaction. *Communication of the ACM*, 42(11), 74-81.
- Oviatt, S. L., Cohen, P., Wu, L., Vergo, J., Duncan, L., Suhm, B., et al. (2000). Designing the user interface for multimodal speech and pen-based gesture applications: state - of - the - art systems and future research directions. *Human - Computer Interaction*, 15, 263-322.
- Pace, S. (2004). A grounded theory of the flow experiences of Web users. *International Journal of Human-Computer Studies*, 60, 327-363.
- Pantic, M., & Rothkrantz, L. (2000). Automatic analysis of facial expressions: The state of the art. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(12), 1424-1445.
- Parkinson, B. (1999). Relations and dissociations between appraisal and emotion ratings of reasonable and unreasonable anger and guilt. *Cognition and Emotion*, 13, 347-385.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods (3 ed)*. Thousand Oaks, California: Sage Publications.

- Peters, R. S. (1970). Concrete principles and the rational passions. In R. S. Peters (Ed.), *Moral Development and Moral Education* (pp. 143-157). London: Allen & Unwin.
- Petrina, S. (2007). *Advanced teaching methods for the technology classroom*: Information Science Publishing.
- Picard, Papert, Bender, W., Blumberg, B., Breazeal, C., Vavallo, D., et al. (2004). Affective learning: A manifesto. *BT Technology Journal*, 22(4).
- Plutchik, R. (1962). *The emotions: Facts, theories & a new model*. New York: Random House.
- Plutchik, R. (1979). *Emotion: A psychoevolutionary synthesis*. New York: Harper and Row.
- Prinz, J. (2004). Which emotions are basic? In P. Cruise & D. Evans (Eds.), *Emotion, Evolution, and Rationality* (pp. 69-87). Oxford: Oxford University Press.
- Ringland, S., & Scahill, F. (2003). Multimodality - The future of the wireless user interface. *BT Technology Journal*, 21(3), 181-191.
- Romaine, S. (1994). *Language in society: An introduction to sociolinguistics*. Oxford: Oxford University Press.
- Roseman, I. J., Antoniou, A. A., & Jose, P. E. (1996). Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition and Emotion*, 10, 241-277.
- Roseman, I. J., Spindel, M. S., & Jose, P. E. (1990). Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology*, 59, 899-915.
- Russell, B. (1960). *An outline of philosophy*. New York: Meridian Books.
- Savage, J. (2000). Ethnography and health care. *British Medical Journal*, 321(7273), 1400-1402.
- Scarantino, A. (2005). *Explicating emotions*. University of Pittsburgh, Pittsburgh.
- Schachter, S. (1957). Pain, fear, and anger in hypertensives and normotensives: A psychophysiologic study. *Psychosom Medical*, 19, 17-29.
- Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. *Advantage Expression in Social Psychology*, 1, 49-80.

- Schachter, S., & Singer, J. (1979). Comments on the Maslach and Marshall-Zimbardo experiments. *Journal of Personality and Social Psychology*, 37, 989-995.
- Schachter, S., & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional states. *Psychological Review*, 69, 379-399.
- Schachter, S., & Singer, J. (1962). Cognitive, social and physiological determinants of emotional states. *Psychological Review*, 69, 379-399.
- Scheffler, I. (1977). In praise of the cognitive emotions. *Teachers College Record*, 79(2), 171-186.
- Scherer, R. (2001). Appraisal considered as a process of multilevel sequential checking. In R. Scherer, A. Schorr & T. Johnstone (Eds.), *Appraisal processes in emotion: theory, methods, research* (pp. 92-129). Oxford, New York: Oxford University Press.
- Searle, J. R. (1980). *Intentionality*. Cambridge, UK: Cambridge University Press.
- Sebe, N., Cohen, I., & Huang, T. (2004). Multimodal emotion recognition. *WSPC for Review Volume*, 16(2), 1-23.
- Seels, & Glasgow. (1990). *Exercises in instructional design*. Columbus OH: Merrill Publishing Company.
- Smith, C. A., & Lazarus, R. S. (1990). Emotion and adaptation. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 609-637). New York: Guilford.
- Solomon, R. C. (1976). *The passions*. New York: Doubleday.
- Solomon, R. C. (1980). Emotions and choice. In A. Rorty (Ed.), *Explaining emotions* (pp. 251-281). Los Angeles: University of California Press.
- Solomon, R. C. (2003). *Not passion's slave: Emotions and choice*. Oxford; New York: Oxford University Press.
- Stocker, M., & Hegeman, E. (1992). *Valuing emotions*. Cambridge: Cambridge University Press.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. New York: Cambridge University Press.

- Suchman, L., & Trigg, R. H. (1986). *A framework for studying research collaboration*. Paper presented at the ACM Conference on Computer-Supported Cooperative Work, Austin, Texas.
- Sully, J. (1902). *An essay on laughter*. London: Longmans.
- Tato, R., Santos, R., Kompe, R., & Pardo, J. M. (2002). *Emotional space improves emotion recognition*. Paper presented at the Proc. of ICSLP - 2002, Denver, Colorado.
- Tomkins, S. S., & Demos, E. V. (1995). *Exploring affect: The selected writings of Silvan S. Tomkins*. New York: Cambridge English.
- Vyzas, E., & Picard, R. W. (1999). *Offline and online recognition of emotion expression from physiological data*. Paper presented at the Third International Conference on Autonomous Agents, Seattle, WA.
- Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching*, 31, 57-71.
- Weizenbaum, J. (1966). ELIZA: A computer program for the study of natural language communication between man and machine. *Communication of the ACM*, 9(1), 35-36.
- Williams, C., & Stevens, K. (1972). Emotions and speech: Some acoustical correlates. *Journal of the Acoustic Society of America*, 52(4), 1238-1250.
- Wolcott, H. F. (1999). *Ethnography: A way of seeing*. Walnut Creek, London, New Delhi: Altamera Pres, A Division of Sage Publications.
- Yob, I. M. (1997). The cognitive emotions and emotional cognitions. *Studies in Philosophy and Education*, 16(1-2), 43-57.
- Zhang, X. (2007). *An ethnographic study on the development of online academic language learning communities in a college-level adjunct ESL program*. The University of Minnesota.



## **Appendix I: Participants Recruitment**

### **Research Study of Technology and Emotion(s) in Chatbots**



## **Chatbot Users Needed**

### **A Study of Chatbot Interface Design For Language Learning**

**Who:**

English as a second language speakers at the intermediate level.

**What and Why:**

A team of researchers at the University of British Columbia is investigating the interface between technology and emotion(s). A commercial chatbot, or digital language tutor that can carry on extensive conversation, will be employed in this research. The research purpose is to explore the nature of users' emotional responses to chatbot technology and add values to the chatbot interface design.

**How:**

Participants will be invited to interact with the chatbot and asked to commit about five hours of their time (one hour per week) to interacting with the chatbot and about two hours to discussing the interactions. Interview comments will be used only with your permission and kept anonymous in any research reports. You will NOT be identified by name in any reports of the completed study.

**Where:**

Five interaction sessions and two interviews will be conducted in a lab. Travel may be necessary.

**Contact:**

For questions or further information please contact:

## Appendix II: Pre-Research Interview

### ***Background Information:***

Name: \_\_\_\_\_ Gender: \_\_\_\_\_ Age: \_\_\_\_\_

Native Language: \_\_\_\_\_ English Level: High Middle Low

### ***Questions:***

1. Did you use computer to learn English before? Yes \_\_\_\_ No \_\_\_\_
2. What are your goals, motives and concerns to participate this research?
3. What kind of learning software did you use to learn English before?
4. How did you feel when using software to learn English via computer?
5. What is your goal to learn English?
6. How did you define yourself as a learner, active learner or passive learner?
7. How comfortable are you in working with computers?  
Not comfortable                  Comfortable                  Very comfortable
8. What do you usually use computers for?

Task	Approx. hours per week used
Games	
Language Learning	
Email	
Working/Writing	
Graphics Design	
Browsing the Internet	
Others	

9. In one sentence, could you summarize how you like the way using computers to learn languages?

## **Appendix III: Interview Questions**

### **Part One: How do you feel when using chatbot Lucy for your language learning?**

1. How would you describe your sense of chatbot Lucy (sensory engagement)?
  - a. In what degree, if any, have you sensed chatbot Lucy: strong sense or weak sense?
  - b. How do you feel your interaction with chatbot Lucy?
  - c. Do you think your involvement and your strong sense of chatbot Lucy will make your learning intrinsically meaningful for you? How? Why?
2. What are your emotional responds to chatbot Lucy?
  - a. Do you believe in chatbot technology can help you learn language effectively? How? Why?
  - b. Which perceived features of chatbot Lucy, do you think, will influence your goals, motive or concerns?
  - c. If there is innovation of chatbot technology, can you imagine which feature will trigger your positive emotions and which feature will trigger your negative emotions?

### **Part Two: How do your emotions serve your cognitive process when using chatbot technology?**

1. What, do you think if any, influences your comfortability in using chatbot technology in your language learning?
  - a. Do you think learning with chatbot Lucy in a public place or in a private place will influence your learning outcomes?
  - b. Do you think your past experience (learning language with computer) will influence your emotions in using chatbot technology in language learning at present?
  - c. Does your past experience in using language learning software will encourage/hinder you to choose new technology such as chatbot in your language learning?
2. How do you define yourself in language learning, for example, a shy student, a quiet student or an active student? What is important for you to learn English?
3. How do you think your relationship with chatbot will influence your emotions in using it?

### **Part Three: What influence your emotional responses to the chatbot technology during learning processes?**



The University of British Columbia  
Office of Research Services  
**Behavioural Research Ethics Board**  
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

## CERTIFICATE OF APPROVAL - MINIMAL RISK

<b>PRINCIPAL INVESTIGATOR:</b> Stephen Petrina	<b>INSTITUTION / DEPARTMENT:</b> UBC/Education/Curriculum Studies	<b>UBC BREB NUMBER:</b> H07-00080
<b>INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:</b>		
<small>Institution</small>	<small>Site</small>	
UBC Other locations where the research will be conducted: UBC room 1224 (Lab space) Home of GRA (Lauren Hall)		
<b>CO-INVESTIGATOR(S):</b> Teresa Dobson Lynn Fels Mary K. Bryson		
<b>SPONSORING AGENCIES:</b> Hampton Research Endowment Fund		
<b>PROJECT TITLE:</b> Encountering Emotion(s) and Technology  *Amendment of B04-0615 (H04-80615)		

**CERTIFICATE EXPIRY DATE:** March 8, 2008

<b>DOCUMENTS INCLUDED IN THIS APPROVAL:</b>	<b>DATE APPROVED:</b> March 8, 2007
<small>Document Name</small>	<small>Version</small> <small>Date</small>
<b><u>Consent Forms:</u></b>	
Main Study Consent	N/A      January 29, 2007
Theatre Study Consent	N/A      January 29, 2007
<b><u>Advertisements:</u></b>	
Chatbots	N/A      January 29, 2007
Robotic Pets	N/A      January 29, 2007
Animators	N/A      January 29, 2007
Theatre	N/A      January 29, 2007
<b><u>Questionnaire, Questionnaire Cover Letter, Tests:</u></b>	
Theatre Questions	N/A      January 29, 2007
Main Study Questions	N/A      January 29, 2007
<b><u>Letter of Initial Contact:</u></b>	
Theatre Contact	N/A      January 29, 2007
Main Study Contact	N/A      January 29, 2007
<p>The application for ethical review and the document(s) listed above have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.</p> <p style="text-align: center;">Approval is issued on behalf of the Behavioural Research Ethics Board and signed electronically by one of the following:</p> <p style="text-align: center;">Dr. Peter Suedfeld, Chair Dr. Jim Rupert, Associate Chair Dr. Arminee Kazanjian, Associate Chair Dr. M. Judith Lynam, Associate Chair</p>	



The University of British Columbia  
Office of Research Services  
**Behavioural Research Ethics Board**  
Suite 102, 6190 Agronomy Road, Vancouver, B.C. V6T 1Z3

## CERTIFICATE OF APPROVAL- MINIMAL RISK RENEWAL

<b>PRINCIPAL INVESTIGATOR:</b> Stephen Petrina	<b>DEPARTMENT:</b> UBC/Education/Curriculum Studies	<b>UBC BREB NUMBER:</b> H07-00080
<b>INSTITUTION(S) WHERE RESEARCH WILL BE CARRIED OUT:</b>		
<b>Institution</b> UBC	<b>Site</b> Vancouver (excludes UBC Hospital)	
Other locations where the research will be conducted: UBC room 1224 (Lab space) Home of GRA (Lauren Hall)		
<b>CO-INVESTIGATOR(S):</b> Teresa Dobson Lynn Fels Mary K. Bryson		
<b>SPONSORING AGENCIES:</b> UBC Hampton Research Endowment Fund - "Encountering Technology and emotion(s)"		
<b>PROJECT TITLE:</b> Encountering Emotion(s) and Technology		
*Amendment of B04-0615 (H04-80615)		

**EXPIRY DATE OF THIS APPROVAL:** April 9, 2009

**APPROVAL DATE:** April 9, 2008

The Annual Renewal for Study have been reviewed and the procedures were found to be acceptable on ethical grounds for research involving human subjects.

Approval is issued on behalf of the Behavioural Research Ethics Board

Dr. M. Judith Lynam, Chair  
Dr. Ken Craig, Chair  
Dr. Jim Rupert, Associate Chair  
Dr. Laurie Ford, Associate Chair  
Dr. Daniel Salhani, Associate Chair  
Dr. Anita Ho, Associate Chair