A Case Study of AI Application on Language Instruction: CSIEC

Jiyou Jia

Department of Educational Technology
School of Education, Peking University
Beijing, 100871. China
jjy@pku.edu.cn

Abstract

CSIEC (Computer Simulation in Educational Communication), is not only an intelligent web-based human-computer dialogue system with natural language for English instruction, but also a learning assessment system for learners and teachers. Its multiple functions including grammar gap filling exercises, talk show, free chatting and chatting on a given topic, can satisfy the various needs from the students with different backgrounds and learning abilities. After the brief explanation of the motivation and the survey of the related works, we illustrate the system structure, and describe its pedagogical functions with the underlying AI techniques in details such as NLP and human-computer interaction. We summarize the free Internet usage from a six months period and its integration into English classes in universities and middle schools. The evaluation findings show that the chatting function has been improved and frequently utilized by the users, and the application of the CSIEC system on English instruction can motivate the learners to practice English and enhance their learning process. Finally, we draw some conclusions for the future improvement.

Task Description

English, as an international language, is treated as a key tool for the development and cultivation of the cross-cultural communicational ability of students, which is gaining more and more attentions in the developing countries like China. English has been listed as one of the obligatory courses in school and higher education. However, some problems exist in the English language education in these countries. First of all, as stressed by situated learning (Brown et al., 1989; Barab and Duffy 2000) and constructivist learning theory (Jonassen, 1994; Von Glasersfeld, 1996), one of the best ways to learn a foreign language is frequent communication with a native speaker. But it is not a practical method in the classroom due to the one-to-one student/teacher ratio it implies. A number of factors ranging from the lack of time to shyness or limited opportunity for quality feedback hamper using the target language (Fryer and Carpenter, 2006).

A potential solution to this problem is to apply computer spoken dialogue systems to role play a conversational partner. If we could design an interactive web-based system which could chat with the English learners anytime anywhere, their great demand for learning partners could be fulfilled. Motivated by the great demand for English instruction, in 2002 we began to design a system which can communicate with the Internet users in English (Jia, 2004, a). Our design principle is application and evaluation oriented. As soon as the system is applicable, we put it into free use in the Internet (http://www.csiec.com) and get the user feedback. We also cooperate with the English teachers from universities and middle schools and integrate the system into English instruction. Through the systematic application and evaluation we get more suggestions and critics, which can direct our research more effectively.

Related Works

Brennan (2006) defined a chatbot as "an artificial construct that is designed to converse with human beings using natural language as input and output". A chatbot architecture integrates a language model and computational algorithms to emulate communication between a human user and a computer using natural language (Abu Shawar and Atwell, 2007).

ELIZA (Weizenbaum, 1966) is the first chatbot. It used key words to analyse input sentence and created its response based on reassembly rules associated with a decomposition of the input. But it held no memory of the conversation. However, the syntactic way of natural language processing exemplified by ELIZA has been developed significantly from 1960s up to now, leading to the development of various chatbots, including ALICEBOT (ALICEBOT). Since 1990s with the improvement of natural language processing, chatbots have become more practical, and have also been applied in education.

Graesser et al. (2005) used “AutoTutor”, an intelligent tutoring system with mixed-initiative dialogue which can simulate a human tutor by holding a conversation with the learner in natural language, to enhance the learner's engagement and the depth of the learning.

Kerfoot et al. (2006) described an experimental use of chatbots as a teaching adjuvant in training medical students.

The experiment showed web-based teaching using chatbots significantly increased test scores in the four topics at each medical school, and learning efficiency was increased three-fold.

Seneff (2006) described several multilingual dialogue systems specifically designed to address the need for language learning and teaching. Several different domains were developed, in which a student's conversational interaction was assisted by a software agent functioning as a tutor providing them with translation assistance at any time.

Abu Shawar and Atwell (2007) developed algorithms for adapting or retraining a chatbot to a training corpus. They stated that the evaluation feedback from language learners and teachers indicated that these adaptive chatbots offered a useful autonomous alternative to traditional classroom-based conversation practice.

Kerly, et al. (2007) discussed the development and capabilities of both conversational agents and open learner modeling. They described an experiment to investigate the feasibility of using a chatbot to support negotiation. The experiment result showed that most students liked the chatbot and the chatbot helped them understand their learner model.

From the related works above we conclude that the usage of chatbot systems in education is drawing more and more attentions from researchers in related fields. This trend confirms our determination to further the development of the CSIEC system and its application in English education.

Current System Architecture

Contrary to the partial parsing used in ELIZA, we attempt the full syntactical and semantic analysis of the user inputs, as the logician G. Frege (1879) pointed out: "The meaning of a sentence exists in the meanings of all words in the sentence and their conjunction method". After parsing the user input text we obtain the user information in the form of XML, i.e. NLML and call them the user facts. The facts are retrieved from natural language expressions, and also represented with the annotation of natural language in the sentence ontology. These facts function as the main contextual source of the robot dialogue reasoning. This thought originates from L. Wittgenstein's theory (1918/21) about the world, facts, objects and human language: "The world consists of facts, the facts consist of objects. The facts are reflected in the language. A logical picture of facts is a thought. The boundary of language is the boundary of knowledge and cognition."

The current CSIEC system is version 9. The whole system consists mainly of the following components, which are illustrated in Figure 1. They are introduced in the sequence of data flow in the program.

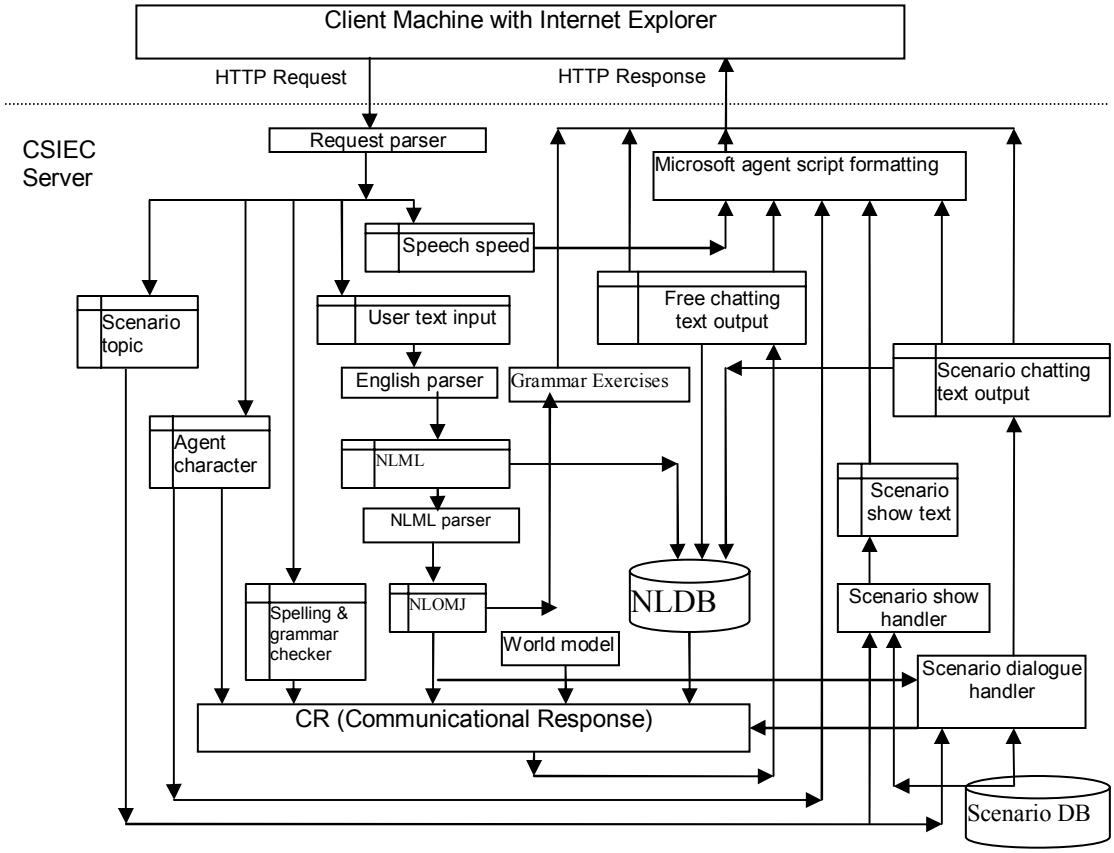

Figure 1: The current architecture of CSIEC system

(1) HTTP request parser resolves the user request from http connection and gets some parameter values: input text, scenario topic, agent character, speech speed, spelling and grammar checker, etc.

(2) English parser parses the user input text into NLML (Natural Language Markup Language). NLML (Jia, 2004, b) is a dependency tree in XML form, and structurally labels the grammar elements (phrases), their relations and other linguistic information in English sentences (words, part-of-speech, entity type, and so on). For example the NLML of the sentence "*I come*" is:

```
<mood>statement</mood>
<complexity>simple</complexity>
<subject>  <noun><type>perspronoun</type>
           <word>I </word><numb>sing</numb>
           <pers>first</pers><case>nom</case>
         </noun>
</subject>
<verb_phrase>
  <verb_type>verb</verb_type>
  <tense>present</tense><numb>sing</numb>
  <pers>first</pers>
  <verb_word>come </verb_word>
</verb_phrase>
```

(3) NLML parser parses the NLML of the user input into NLOMJ (Natural Language Object Model in Java) which represents the grammatical elements and their dependency with the Sentence ontology in the working memory (Jia, et al., 2004). Through NLOMJ the declarative sentence is retrieved and decomposed into atomic facts consisting of only one subject and one verb phrase.

(4) NLDB (Natural Language Database) stores the historical discourse, the user atomic facts in the form NLML, the robot atomic facts which are also expressed in NLML, and other data.

(5) World model contains common sense knowledge which is the basis for response generation and logical inference. It is now represented by WordNet (Fellbaum, 1998).

(6) CR (Communicational Response) mechanism comprehensively takes into accounts the user input, the user facts stored in NLDB, the world model, the personality of the user expressed in the previous dialogue, and that of the robot itself selected by the user. The response is generated by this special mechanism. We illustrate it with the comparison of two human-computer dialogues with the ELIZA-like ALICEBOT and with our system. The dialogue with ALICEBOT is:

Human: hello
ALICE: Hi there!
Human: I am Peter
ALICE: I could do it more if I wanted to
Human: who am I?
ALICE: You are unknown person, seeker.

The dialogue with CSIEC is:

You: hello!
CSIEC: hi, what is your name?
You: I am Peter.
CSIEC: hello, Peter.
You: who am I?
CSIEC: You are peter.

(7) Scenario dialogue handler creates the robot output corresponding to the user input within a given scenario.

(8) Scenario show handler creates the random robot-robot talk show scripts within a given scenario.

(9) Scenario DB stores the robot-robot talk show scripts and human-robot dialogue scripts which are manually written by designer, for example English language teacher.

(10) Microsoft agent script formatting transforms the output text into VB script, considering the selected agent character and speaking speed.

(11) Browser/Server interface processes the HTTP request from client machine and responds with the robot output, either in text or with VB script.

The system is implemented in JDK1.6 and uses MySQL as database.

Basic Functions and Underlying AI Technologies

Among the various functions of the CSIEC system we mainly introduce its specific chatting function.

Multimodal User Interface and Selectable Chatting Pattern

Like human being's authentic dialogue situation, the Internet users have different preferences for the dialogue simulation. In order to adapt to variant user preferences the CSIEC provides several user interfaces and dialogue patterns.

At first the users can chat with the robot either through text or via voice. They can hear synthesized voice and watch the avatar performance through Microsoft agent technology. They can speak to the robot through a microphone with a speech recognition program like IBM ViaVoice.

Secondly the robot can check the spelling and grammar of the input upon the user's request.

Thirdly the chat topic between the user and the robot can be either free (unlimited) or given (limited). The unlimited dialogue simulation doesn't specify the dialogue topic and content. It benefits users whose English is fluent or who are at least good at written English, as well as users who are extroversive or conversational. However, users whose English is poor, or who are introversive, have little to chat with the virtual chatting partner. For them an instructive dialogue guided by the agent is more helpful. Language teachers also acknowledge that conversation practice is normally on a specific topic during the learning of topic-specific vocabulary and language.

It is noticeable that in normal human talking these two chatting patterns are not absolutely separated, but often

interleave each other. In our system design we consider this interaction too.

Free Chatting Adaptive to User Preference and Topic

In the free chatting the users with different characters and personalities may choose different types of chatting pattern. For the sake of user dialogue personalization we have designed five Microsoft agent characters which represent different kinds of chatting patterns. Christine always tells the user stories, jokes and world news. Stephan prefers to listen quietly when the users share with him their own experiences. Emina is a curious girl, and is fond of asking users all kinds of questions related with the users' input. Christopher provides comments, suggestions and advices on the user's input. Ingrid behaves as a comprehensive virtual chatting partner, who gives users responses considering both the input text and the discourse context.

Upon user registration to the chatting system the user's profile is obtained and recorded, such as the gender, birthday, educational level, address. So the corresponding chatting topic and content can be generated based on the personal information. Of course if the user expresses the wish to change the chatting topic during the process of the robot's narrating comments or asking questions, the robot should terminate this process and transfer to another topic given by the user or by the robot itself. If the user specifies a topic, for example, "*I want to talk about sport*", the robot changes the topic to it. If the user just expresses the wish to change the topic, but does not determine a topic, such as "*I want to talk about another topic*", the robot selects one from the waiting topics list which has not been talked about.

The user's interests are also expressed in the input, e.g. through the mentioned nouns and verbs. So the chatting topic can be triggered by nouns and verbs. More frequently one noun is or several related nouns are talked about, the related topic is more emphasized. Thus the chatting between the user and robot can be regarded as guided chatting or chatting on the topic.

We deal with the chatting on a given topic in two ways. One is predefining some comments or asking some questions about this topic. By talking about it only one statement or question will be randomly selected and given out. Another is to search the topic in the guided chatting within a given scenario, and then transfer the chatting to the guided chatting in a given scenario, which will be introduced more in next subsection. The arrow from the scenario dialogue handler to the communicational response in Figure 1 indicates this relation.

Summarily the goal of free chatting is to motivate the user's talking desire. For this purpose the robot tries to adapt itself to the user's interest, and launch new topics.

Guided Chatting in a Given Scenario

The dialogue on a given topic requires not only the common sense knowledge, but also the domain knowledge, in addition to the explicit usage of the knowledge with the form of natural language. The dialogue should be developed step by step around a red line or a topic for this scenario. It can be described by a decision tree structure with many branches. We use scripts to describe the decision tree in the dialogue on a given topic. The script is made up of lines of dialogue states, every of which is a branch in the decision tree. Suppose the robot speaks at first. In every line there must be the text output from the robot and its order number in the dialogue. This output may be triggered by specific user input, which we call the prerequisite of this output text. The robot may also expect the user inputs certain texts, or some texts with specific semantic or syntactical characters, which we call the expectation of this output text. We write the line in the script with the format:

Nr. <prerequisite> (text) <expectation>.

The "Nr." and "text" are the necessary two components in every line. The "Nr." is an integer indicating the line order in the whole script, whereas the "text" can be any text from the robot, either statement, or question, and so on, and it is written within closed brackets.

In a script line the prerequisite and expectation are optional. If they appear they must be written within closed sharp brackets. If the prerequisite exists and is satisfied the output text is given out by the robot. The expectation means the robot hopes the user responses to this text with some specific syntactic and/or semantic features, and can be applied to instructional goal. For example if the user's input does not satisfy the robot's expectation he/she will face the previous robot output again, until the expectation is fulfilled. This dialogue pattern can be used for drill. Another alternative is that the user is given a high mark if the input satisfies the robot output, otherwise a low mark, although the robot continues the next dialogue. This pattern can be used in test or examination.

The format of the prerequisite is:

<Nr, variable 1: value 1, value 2...; variable 2: value 1, value 2...>

The format of the expectation is:

<variable 1: value 1, value 2...; variable 2: value 1, value 2...>

The prerequisite needs an order number indicating the expectation of which line this condition fulfills. There may be more than one value for a given variable. This means if the variable equals any one of the listed values, the condition is fulfilled, i.e. the values for a given variable have the relation of logical disjunction. There may be also more than one variable and its corresponding values. The relation among these variables is the logical conjunction.

One example script is about the "*salesman and costumer*". Its script texts are:

1. (Can I help you?)<coat<hyponym><n>:coat>

1. (May I help you?)<coat<hyponym><n>:coat>

2.<1,coat<n>:coat>(What color do you like?)<color <hyponym><n>:white,any>

3.<1, coat<n>:>(We have only coats. Good bye.) <terminator>

4.<2,color<n>:white>(How about this one?) <price:price>

5.<2,color<n>:any>(How about this white one?) <price:price>

6.<2,color<n>:>(We have only white ones. Bye bye.)<terminator>

7.<price:price>(It costs only twenty Yuan.) <response <j>:cheap,expensive>

8.<price:> (But it is very cheap. It costs twenty Yuan.) <response<j>:cheap,expensive>

9. <response<j>:cheap>(Thanks.)

10.<response<j>:expensive>(Sorry. That's the best we can do.)

11. <response<j>:>(Thanks. Bye bye!)

In the prerequisites and expectations of this example there are some new symbols.

<hyponym>: the user's input satisfies the expectation, if it contains a hyponym of the values. For example: for the expectation "<coat<hyponym> <n>:coat>", the user's input "I want to buy a cutaway" fulfills it because cutaway is a kind of coat.

<n>: the variable and its values are nouns, and the user's input satisfies the expectation, if it contains a synonym of the values.

<j>: the variable and its values are adjectives, and the user's input satisfies the expectation, if it contains a synonym of the values.

<d>: the variable and its values are adverbs, and the user's input satisfies the expectation, if it contains a synonym of the values.

<terminator>: this is the last output from the robot, and the discourse for this scenario ends after it and the user's next response, whatever it may be.

The lines 7 to 11 demand a prerequisite without any line number. So they can appear after any user input which satisfies the condition and has nothing to do with the line order. The most lines in this script depend on the preceding ones one by one, so that this script corresponds to a deeper tree structure.

We use the keyword detection plus semantic analysis to interpret these scripts. The script interpreter handles the variable values in the prerequisite and expectation as the keywords (and their synonyms, hyponyms, etc.) in the user input within the framework of syntactical and semantic analysis.

Nevertheless, this kind of script is difficult to be written by human, for example the English teachers who want to use this program to train the students. Thus we have designed a Java GUI, i.e. DSE (Discourse Script Editor) for editing the scripts step by step more easily (Jia and Chen, 2008).

Automatic Scoring of Gap-filling Exercises without Defined Answers

Traditional computer-based gap filling exercises require a definite answer or a set of definite answers. For the questions whose answers are difficultly to be listed, the human manual check is still unavoidable. However, this kind of exercise without predefined answers can promote the creative thinking of the students.

With the spelling and grammatical check function the CSIEC system can decide if a filled gap-filling sentence is grammatically correct. Therefore it can be applied to assess the gap-filling exercises and relieve the teachers' burden. So currently the system provides the interface for teachers to design new gap-filling exercises, as well as the interface for learners to do these exercises and then to get the automatic assessment results.

An example of gap-filling exercises is: "*I ( ) a student.*" The correct answer to the gap can be: "*am*", "*want to be*", "*will be*", "*have been*", "*need*", "*help*", etc.

Talk Show of Two Robots

This function is designed to aid the user's chatting with the robot on a given topic. With it the users can watch the talk show of two robots before the human-computer interaction. The talking texts are predefined by the teacher for the specific context or topic. However, the actual texts for a given meaning can be expressed randomly. So this kind of talk show is different from the monotone one presented in the traditional video or audio cassette. It will enforce the learner's spontaneous listening and understanding.

The talk show script texts can be readily written by the teachers with any text editor.

Listening Training

We use the Microsoft agent technology to synthesize the output text, because the agent's voice is lifelike, the agent's figures, movements as well as actions can be designed very vividly, and it can synchronously display the spoken text, which facilitates the aural understanding and activates the user's interests. We have also designed seven facial expressions (neutral, happy, sad, feared, disgusted, angry and surprised) for every agent character. The robot's reading speed can be adjusted by the users at any time. We have also designed a free webpage whose agent can read any texts inputted or pasted by the user.

Different from the traditional audio technologies such as audio players, the user confronts with unexpected robot text and voices, like talking with a real human being. So this function can benefit the user's listening comprehension and prompt response.

Scoring Mechanism

In order to motivate the users to learn English we trace users' usage of different functions and give them certain scores. The underpinning score principle is encouraging the usage of chatting with agents, and with spelling and grammar checking. By the chatting on a given context, the user is given a high mark if the input satisfies the robot output, otherwise a low mark. This mark also contributes to the total score.

The user can review his performance and scores after entering the system. This function is very important and helpful for self learning and evaluation. A special user who

is labeled as the teacher can access the performance and scores of all the users who are classified as his/her students. This automatic monitoring function is very necessary for the teacher to assess the students' learning behavior and progress.

Application and Payoff

Summative Evaluation of Free Using in Internet

The internet users come to the CSIEC webpage mainly through search engines, because our website has become one of the top 5s in the searching results of famous search engines such as google.com, yahoo.com and baidu.com by related keywords such as "English chatbot", "Online English learning" in Chinese or in English, although we haven't made any large-scale advertisement. The effectiveness and attractiveness of the system have been demonstrated by the achievement.

With the recorded human-computer dialogues, we can summarize the system's chatting function from Jan. 20th 2007 to June 20th 2007. The users with different user names who accessed the CSIEC during this period count 1783.

The chatting quality can be measured by the chatting duration between the user and the robot, as defined in Jia (2004, a)'s work on the experimental report of using ALICEBOT for English learning. To calculate the chatting duration we define two terms: round and number of the rounds. A round means a user input and a corresponding robot output to the user. The total rounds of a given user cover all dialogs between the user and the chatbot, and can be used to measure the duration of the user's chatting with this chatbot.

Sometimes the user cannot get any response from the chatbot after a long waiting time. The reasons may be the multi-users' simultaneous visits to the server, and some bugs in the system design. After scrutinizing the user dialogue records we find 313 such error rounds, and 48840 effective rounds with both input and output. We define the natural error rate as the ratio of the error rounds number divided by the total rounds number. In this study period the natural error rate is 313/(313+48840), i.e. 0.64%.

The average rounds number is 27.4. The number of the rounds varies from 1 to 580. 48.85% of the users chat with the robot briefly (<=10 rounds); 38.42% chat with it long (between 10 and 50 rounds); 7.63% chat with it longer (between 50 and 100 rounds); and only few, 5.1%, chat with it very long (>100 rounds). Compared with the finding in (Jia, 2004, c), the percentage of the brief chatting with the robot has decreased by 21.78%. Proportionally the percentage of the long and longer chatting has increased.

User Feedbacks

In the foot of almost every webpage of the CSIEC system we leave a feedback text area so that the users can straightforwardly enter their comments, critics and suggestions either in Chinese or in English. Through the detailed securitizing and analysis of these texts, we hope to find what problems are addressed by the users.

Many users still input normal chatting text into this area such as "*hello*", and other nonsense texts. Excluding them we get 341 lines of real feedback.

There are 79 positive comments. 37 of them are very simple positive comments such as "*(very) good*", "*(very) well*". Six emotionally praise the system without any reason, for example: "*clever*", "*I like this*", "*I love you*". 36 express positive comments with reasons, e. g.: "*The robot is more advanced than before, and also personalized.*", "*The access speed is faster than before.*", "*The kind of communication can improve our English.*"

Seven feedbacks are very simple negative comments without any reason, such as "*not good*", "*just so so*", "*simple*", "*stupid*". Others pointed out the problems they met: they can neither see the agent animation nor hear the agent's speech, the agent character is not good, the agent voice sounds weird, the agent speaks too fast, or the agent speaks too slowly (12); can't understand what the robot is saying, or the meaning of new words (13); there are grammatical errors in the robot's responses (5); the dialogue in a given scenario is too short (10); the robot only asks the user a same question, but can't answer the user's question or the answer is false (15); the robot talks little, or often changes the topic (12); the response from the robot is (too) slow (16).

Moreover, ten of the 341 feedbacks give both positive and negative comment, like "the system is very good, but the robot responds too slowly."

The feedbacks point out either technical problems or content shortcomings, which should be overcome in the further improvement.

Formative Evaluation of English Class Integration

After discussing with the English teachers about the class integration and evaluation of the CSIEC system we decided that the instructional instruments are the talk show by two chatting robots with the scenario content and the student' talking with one robot for a given topic corresponding to the textbook content.

In the first term 86 graduate students from two English classes taught by the same teacher participated in the study, and the teacher only recommended the students to use the system, but didn't require them. For the 12 teaching units we designed 25 scenario scripts of human-robot chatting. In the second term 45 high school students in Grade 2 attended the study, where the teacher required the students to use the system together in the computer room. For the 10 teaching units we designed 40 scenario scripts.

We collected data from questionnaires completed at the end of the experimental term. Because in the high school the learning of the content in every course unit is stressed by the teacher, an item of "Reviewing key points in the course units" is specially added in the questionnaire. All the items were measured with a five-points Likert

agreement scale, i.e. the value 5 indicates the maximum best agreement, and 1 means no agreement. For the 5 items of the graduate students the Cronbach's Alpha is 0.933. So the reliability of these surveyed items is very good. For the 6 items of the high school students the Cronbach's Alpha is 0.741. The reliability of these surveyed items is not very good, but acceptable.

We can't recognize big difference between the graduate students' attitude and that of the high school students. It shows the students feel the CSIEC-based English learning can help with course unit review, make them more confident, improve their listening ability, and enhance the interest in language learning. Another item in the questionnaire for high school students shows 60.5% of the students "*like*" or "*like very much*" such a form of English learning, whereas only 2.3% dislike it. 60.5% of them will continue using the system after class, even without the teacher's request.

In third term the system is integrated into an English class in Grade one of junior middle school. The comparison of two examination results before and after the integration class shows great improvement of students' performance, and the survey data also indicates the students' favor to this system. For example the average exam score of the whole class was improved from 64.39 to 90.81, whereas the standard deviation is decreased from 20.129 to 9.572. All the students hope to continue to use this system in English learning. Figure 2 shows two photos taken in the middle school computer rooms using the CSIEC system.

Figure 2 Middle school students using CSIEC in computer classroom

Application Development and Deployment, and Maintenance

This application driven research project has been developed since 2002, and freely used by the Internet users since its birth. Up to now the registered users count more than 30,000, and the first webpage is clicked more than 500 times every day.

We also cooperate with the English teachers and integrate the system into English instruction. Since 2006 it has been used by four university classes, three middle school classes. Together more than 500 students have used it in English class.

Our project has been sponsored by the Konrad-Adenauer-Foundation Germany, Ministry of Education China, Capital Beijing, Peking University and Korea Foundation for Advanced Studies. It won the second prize in the national "*Innovation Competition*" in 2006 held by Ministry of Education and Ministry of Science and Technology, China. With these supports the CSIEC research group has been built since 2006. It includes several graduate and undergraduate students, and the leader is still Dr. Jiyou Jia, the founder. The system is now located in the campus network of Peking University, maintained and updated by Dr. Jiyou Jia.

Discussion and Conclusion

The original goal of the system is supplying a virtual chatting partner for English learner. So the chatting is the most fundamental function. The statistical analysis about the users' behavior indicates that the users have a preference for chatting without spelling and grammar checking. This fact proves that the users prefer the unique chatting function which is lacked in other systems. We must continue to reinforce this primary utility.

The chatting quality can be somewhat shown by the chatting length. The increased percentage of the long and longer chatting shows that the free chatting quality of CSIEC is getting better. The underlying design principle, i.e. fully syntactical and semantic analysis of the user input, and communicative response mechanism, as well as the effort of chatting personalization and adaptation contribute to the chatting quality progress.

The chatting on a given topic is mainly used by the students in the evaluation study, and is also the main function of the whole system the students have used. The formal evaluation results indicate the application of CSIEC system in English class can better assist their language learning, e.g., the confidence on English communications, and the interest in learning English, help them master practical expressions, and improve listening skills. The planned system functions including free chatting and chatting on a given topic, and listening training have been brought into actual pedagogical play.

Through the application and evaluation we find currently there are still some user requirements which haven't been fulfilled well, like the system's stronger ability of natural language understating and generation, which is the fatal factor influencing the human-computer communication, the lifelike synthesized agent voice and high response speed, which also have been addressed in the users' feedback. Solely in NLP many problems are still hard to be solved, such as the textual ambiguity and entailment, which are critical to natural language understanding and generation capability of the CSIEC system, too. How to tackle these problems is a great challenge to us.

As for the system application and evaluation, we will continue the cooperation with the English teachers and monitoring the English ability change of the students' using the CSIEC system. On the other hand we will explore more powerful applications of the underlying techniques of this system in other related fields, such as computer aided test for language learning, computer assisted writing and translation, etc.

References

Abu Shawar, B., and Atwell, E. 2007. Fostering language learner autonomy via adaptive conversation. In: *Proceedings of Corpus Linguistics 2007*.

ALICEBOT. URL: http://www.alicebot.org.

Brown, J., Collins, A., and Duguid, P. 1989. Situated Cognition and the Culture of Learning. *Educational Researcher*, 18(1): 32-42.

Fellbaum, C. 1998. *WordNet: An Electronic Lexical Database*. MA: MIT Press.

Frege, G. 1879. *Begriffsschrift, eine der arithmetischen nachgebildete Formalsprache des reinen Denkens.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Fryer, L., and Carpenter, R. 2006. Emerging technologies bots as language learning tools. *Language Learning &Technology,* 10 (3): 8-14.

Graesser, A. C., Chipman, P., Haynes, B. C., and Olney, A. 2005. AutoTutor: An intelligent tutoring system with mixed-initiative dialogue. *IEEE Transactions on Education,* 48 (4): 612-618.

Honebein, P. 1996. Seven goals for the design of constructivist learning environments. In Wilson, B. (Ed.): *Constructivist learning environments: Case studies in instructional design*, 11-24. NJ: Educational Technology Publications.

Jia, J. 2004, a. CSIEC (Computer Simulator in Educational Communication): A Virtual Context-Adaptive Chatting Partner for Foreign Language Learners. In Kinshuk, et al. (Eds): *Proceedings of ICALT04 (International Conference on Advanced Learning Technology)*, 690-692. IEEE Press.

Jia, J. 2004, b. NLML - a Markup Language to Describe the Unlimited English Grammar. CoRR cs.CL/0404018.

Jia, J. 2004, c. The Study of the Application of a Web-Based Chatbot System on the Teaching of Foreign Languages. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology & Teacher Education International Conference 2004*, 1201-1207. VA: AACE.

Jia, J., and Chen, W. 2008. Script Based Design for Human-Computer Dialog in Given Scenarios for English Learners. In: *Proceedings of ICALT08 (International Conference on Advanced Learning Technology by IEEE Computer Society, Spain)*, Forthcoming.

Jia, J., Ye, Y., and Mainzer, K. 2004. NLOMJ--Natural Language Object Model in Java. In Z. Shi, et al. (Eds.): *Intelligent Information Processing II. Proceedings of ICICP04*, 201-209. Springer.

Jia, J., Chen, W., and Hou, S. 2006. Improving the CSIEC Project through Agent Technology & Adapting It to English Instruction in China. In: *Proceedings of International Conference of Computer on Education*, 47-54. Beijing Normal University.

Jonassen, D. H. 1994. Thinking Technology: Toward a constructivist design model. *Educational Technology*. 1994, (3): 34-37.

Kaufman, D. 2004. Constructivist issues in language learning and teaching. *Annual Review of Applied Linguistics*, 24: 303-319.

Kerfoot, P., et al. 2006. A Multi-institutional randomized controlled trial of web-based teaching to medical students. *Academic Medicine,* 81, (3): 224-230.

Kerly, A., Hall, P., and Bull, S. 2007. Bringing chatbots into education: Towards natural language negotiation of open learner models. *Knowledge-Based Systems*, 20(2): 177-185.

Seneff, S. 2006. Interactive Computer Aids for Acquiring Proficiency in Mandarin. In: Huo et al. (Eds.): *Proceedings of ISCSLP 2006*, *LNAI 4274*, 1–12. Springer.

Von Glasersfeld, E. 1996. Introduction: Aspects of constructivism. In Fosnot, C. (Ed.): *Constructivism: Theory, Perspectives, and Practice*, 3-7. NY: Teachers College Press.

Weizenbaum, J. 1966. ELIZA – a computer program for the study of natural language communications between man & machine. *Communications of the ACM,* 9(1): 36–45.

Wittgenstein, L. 1918/21. *Tractatus logico-philosophicus.* Frankfurt am Main: Suhrkamp.