Spoken Dialogue Systems for Language Learning[*]

Stephanie Seneff, Chao Wang, and Chih-yu Chao
Spoken Language Systems Group
MIT Computer Science and Artificial Intelligence Laboratory
The Stata Center, 32 Vassar Street, Cambridge, MA 02139, USA
{seneff,wangc,chihyu}@csail.mit.edu

Abstract

This demonstration will illustrate interactive computer games intended to help a native speaker of English learn Mandarin. These systems provide users with human-like conversational exercises with contextualized help mechanisms. Two distinctly different activities, a *translation game* and a *dialogue game* are illustrated. The level of difficulty can be manipulated, and the sentence variations covered by the systems familiarize users with different expressions of the same meaning. The systems preserve the qualities of a typical computer system, being infinitely patient and available any time of day. Students will be able to repeatedly practice conversation with no embarrassment.

1 Introduction

Mandarin Chinese is one of the most difficult languages for a native English speaker to learn. Chinese is substantially more difficult to master than the traditional European languages currently being taught in America – French, Spanish, German, etc., because of the lack of common roots in the vocabulary, the novel tonal and writing systems, and the distinctly different syntactic structure.

It is widely agreed among educators that the best way to learn to speak a foreign language is to engage in natural conversation with a native speaker of the language. Yet this is also one of the most costly ways to teach a language, due to the inherently one-to-one student-teacher ratio that it implies.

Recent research in the Spoken Language Systems group at MIT has focused on the idea of designing entertaining computer games as a device for teaching a foreign language, with initial emphasis on the language pair, English and Mandarin. The games are accessible at a Web page, and the student's speech is captured from a headset microphone to support natural spoken dialogue interaction. The system can also be installed to run completely stand-alone on the local laptop computer.

2 Demonstrated Systems

The demonstrated systems comprise two related activities, the *translation game* and the *dialogue game*. The translation game serves as preparation for the dialogue game: the user acquires expertise in speaking within the domain in the target language. The system randomly presents sentences in English and asks the student to speak a sentence of equivalent meaning in Mandarin. To imitate the competitive spirit of video games, the system offers ten *difficulty levels,* which are automatically adjusted depending on the student's monitored performance. After advancing to the highest difficulty level, they will subsequently be much better equipped to converse with the system within the dialogue game.

The *dialogue game* involves spoken conversational interaction to solve a particular scenario. The student and computer are tasked with jointly solving a specified goal. Differing difficulty levels are achieved via the device of a robotic tutor who assists the student in solving their side of the conversation.

2.1 Translation Game

The translation game is motivated by the learning approach advocated by Pimsleur (1967). By practicing translation repeatedly, language learners are

---

able to internalize the structures of the target language, and thus the vocabulary, grammar rules, and pronunciation are practiced concurrently. The user begins by translating isolated vocabulary items in Level 1, advancing to phrases and full sentences at higher levels. The most difficult level, Level 10, involves long and complicated sentences.

We have implemented this game in two domains: (1) flight reservations, and (2) hobbies and schedules. Details of the translation procedure can be found in (Wang and Seneff, 2006), and the algorithm for assessment is described in detail in (Wang and Seneff, 2006). The input utterance is processed through the speech recognizer and language understanding (Seneff, 1992) components, to achieve a simple encoding of its meaning. The system compares this meaning representation to one automatically derived from the targeted English equivalent. The system then speaks a paraphrase of the user's hypothesized utterance in both Chinese and English (Baptist and Seneff, 2000). If it has determined that the student was successful, it congratulates them and prompts them with the next English sentence for translation. At any time, the student can ask for assistance, in which case the system will provide them with a "correct" translation of the English utterance, which they can then attempt to imitate.

2.2 Dialogue Game

In the dialogue game (Seneff, 2006), the user is asked to solve a particular scenario, by role playing a specified persona, which changes dynamically every time the game is played. We will demonstrate the dialogue game in the hobbies and schedules domain. The student is provided with a specification of their preferences for participating in possible activities (swimming, dancing, watching movies, etc.) as well as a calendar specifying activities they are planning to do in the next few days. They are tasked with arranging with the computer to jointly participate in an activity that they both like, at a time when both are free. Another option is for either party to invite the other one to join them in an activity that is already on their schedule.

In addition to the robotic dialogue partner, the student is assisted in solving the task by a robotic *tutor*, who helps them plan what to say next. The tutor works with the same information that the student has, and independently plans the student's half of the conversation. At each dialogue turn, it provides a proposed response, based on the evolving dialogue context. Five different difficulty levels have been implemented, as follows:

1. *Eavesdropping:* The student can simply let the tutor carry out their side of the conversation by clicking a button to advance each dialogue turn.
2. *Parroting:* The system presents a proposed sentence in pinyin on the screen, and the student can just read it out loud well enough to be successfully understood.
3. *Translation:* The system presents an English sentence which the student needs to translate into Chinese.
4. *Characters:* The system presents the Chinese sentence in a character encoding.
5. *Solo:* The tutor stops being pro-active, but can be consulted if necessary.

Both the translation game and the dialogue game will be illustrated live in the demonstration. The systems can be evaluated by two types of basic performance measures: (1) for each system, the recognition accuracy and the translation accuracy serve as an index of quality; (2) calculating the success rate in the translation game and the number of turns taken to complete each dialogue will provide a quantitative view of interaction. Also a pre- & post-test design in the user study will further confirm the pedagogic value of the systems. Ongoing and future work involves expanding the domains supported and introducing the games to the classroom setting.

References

Baptist, L. and S. Seneff. 2000. "Genesis-II: A Versatile System for Language Generation in Conversational System Applications," *Proc. ICSLP*, III:271–274.

Pimsleur, P. 1967. "A Memory Schedule," *Modern Language Journal*, 51:73–75.

Seneff, S. 1992. "TINA: A Natural Language System for Spoken Language Applications," *Computational Linguistics*, 18(1):61–86.

Seneff, S. 2006. "Interactive Computer Aids for Acquiring Proficiency in Mandarin," Keynote Speech, *Proc. ISCSLP*, pp. 1–11.

Wang, C and S. Seneff. 2006. "High-quality Speech Translation in the Flight Domain," *Proc. INTERSPEECH*.

Wang, C. and S. Seneff 2007. "Automatic Assessment of Student Translations for Foreign Language Tutoring," *Proc. NAACL-HLT*.