ZIP & TERRY: A new attempt at designing language learning simulation

Rong-Chang Li
David Topolewski
English Xchange, Inc.

ZIP & TERRY is a simulation designed for children to learn English as a second language (ESL). This article introduces the contents and structure of the program, discusses issues such as balance of entertainment and education, and uses of speech recognition technology in language learning simulation.

Computer games and simulations have long been used for language learning. A review of more than 60 research studies revealed that games and simulations were more effective than traditional classroom instruction, and the effectiveness was strongest for language arts (Randel, Morris, Wetzel, & Whitehill, 1992). The newly produced English as a second language (ESL) simulation ZIP & TERRY is an English learning program for children marketed in China by C-Interchange, Inc. With the application of automatic speech recognition (ASR) technology and a design by experts from the entertainment industry, natural language understanding (NLU), neurolinguistics, learning methodologies, and vocabulary acquisition, the program provides an English learning environment where children are supposed to learn to speak English as they do in a real English-speaking world.

Story

The story is about a young alien named Zip, who comes from the planet Topo. He is given a homework assignment to study a newly discovered language—English. While circling Earth, his spaceship crashes into the Broccoli family home. To survive on Earth, Zip has to communicate with the Broccoli family and learn English. As the learner, you must act as Zip, who must repair his spaceship to get back home. Repair parts appear once you have successfully completed certain learning activities. The learner's progress in English moves him or her toward the learning goal step by step. Zip also participates in the Cosmic Challenge Quiz Show, whose Grand Prize is a trip to Astronaut Space Camp where Zip can meet outstanding scientists who will help to repair the spaceship.
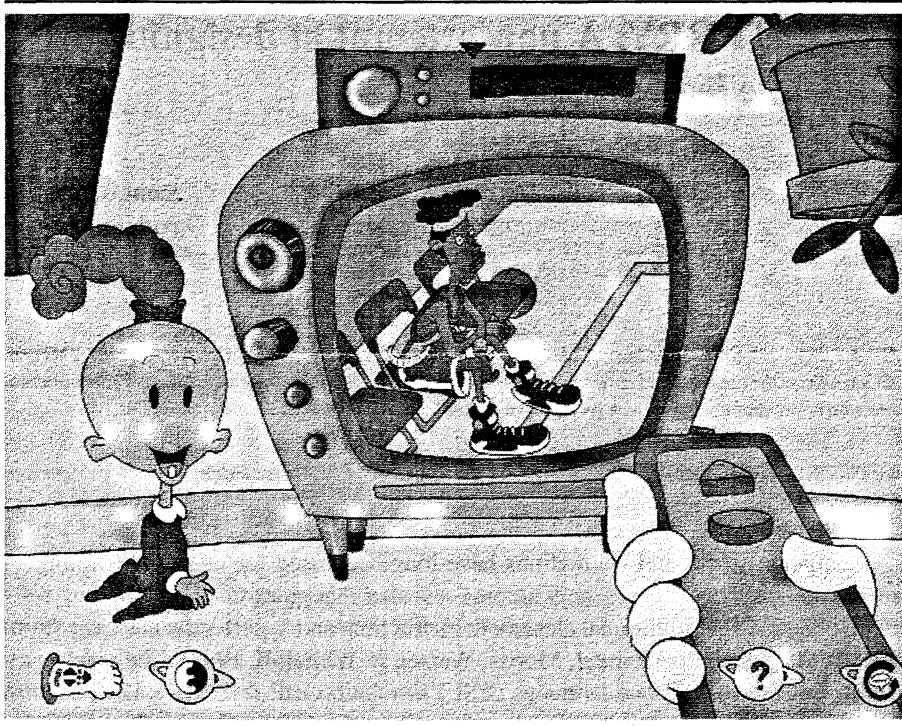
FIGURE 1:   Cindy in the Living Room Environment

Environments

Several realistic environments were designed for the English learning activities—Living Room, Kitchen, Yard Sale, Terry's Room, and Lucky's Corner—where the learner interacts with various characters such as Mr. and Mrs. Broccoli, Terry (the son), Cindy (the baby daughter), and Lucky (the pet pig). Activities include conversations with the characters, playing card games to help to memorize nouns and pronouns, learning verbs by surfing through television channels, taking on the job as salesperson at the Yard Sale, and giving commands to the pet pig, who responds with silly pet tricks. Figure 1 shows how the screen appears when the learner is surfing channels.

There are four modes for each of the environments of Living Room, Kitchen, and Yard Sale. When the program is first used, the learner must proceed in the order that each mode is assigned. After successfully completing all, the learner may work in any particular mode of his or her choice.

1. Vocabulary Development Mode (VDM). VDM is where the learner is challenged with a series of vocabulary training exercises related to the selected environment.

VDM is designed to optimize the process of vocabulary learning and has 12 different features to help the learner whenever it is needed.

2. Environment Research Mode. This is one that reviews the words learned in VDM in a fun, interactive way. An environment appears without any objects in it. The learner places objects in it by saying the name of each object correctly. ASR is used to judge if the learner has pronounced the word well enough. There is also a row of cookies placed in front of the pet pig. The more correct objects placed in the environment, the more cookies Lucky eats. If more than 25% of the objects cannot be identified correctly, the learner must repeat VDM.

3. Conversation Training Mode. This mode is where the learner interacts with people and objects in the environment. A host character asks the learner to listen and repeat after her so that commonly used phrases and patterns are acquired. The characters respond verbally with short action sequences.

4. Activity Mode. This features two types of activities. In the first activity, the learner must listen carefully, comprehend what he or she is asked to do, and is rewarded with congratulations and entertaining animation if successfully accomplishing a task. The second activity is answering questions orally. Characters in different environments ask you *yes/no, either . . . or,* and *how* and *what* questions. The speech recognizer will judge if verbal answers are correct, and give feedback and encouragement.

In the Terry's Room environment, the learner chooses phrases from the Phrase Book to talk to members of the Broccoli family. When the learner asks one of the questions from the Phrase Book, he or she will get an interesting answer from the character on the screen. Lucky's Corner (see Figure 2) is an environment where the learner gives orders and the pig Lucky will act accordingly. If the learner says, "Jump!" the pig will jump, and if the learner says, "Sing!" the pig will sing a song.

Features of ZIP & TERRY

One of the principles we tried to maintain in the design of the program is learning by doing (Schank & Cleary, 1995). A simulation program creates a foreign language environment in which learning is meaningful. In Terry's Room, if Terry is asked what his hobby is, he will say he likes writing e-mail. If the learner asks the pet pig to jump, the pig will jump. If Grandpa is asked to pass the salt on the table, the learner gets the salt. The language is used to do certain things and the results are visible on the screen. A simulation program plus ASR technology makes it possible to create such a language learning environment.

Motivation is important in any language learning. In ZIP & TERRY, the learner has an overall goal—helping Zip return home. In addition to rewards for every learning
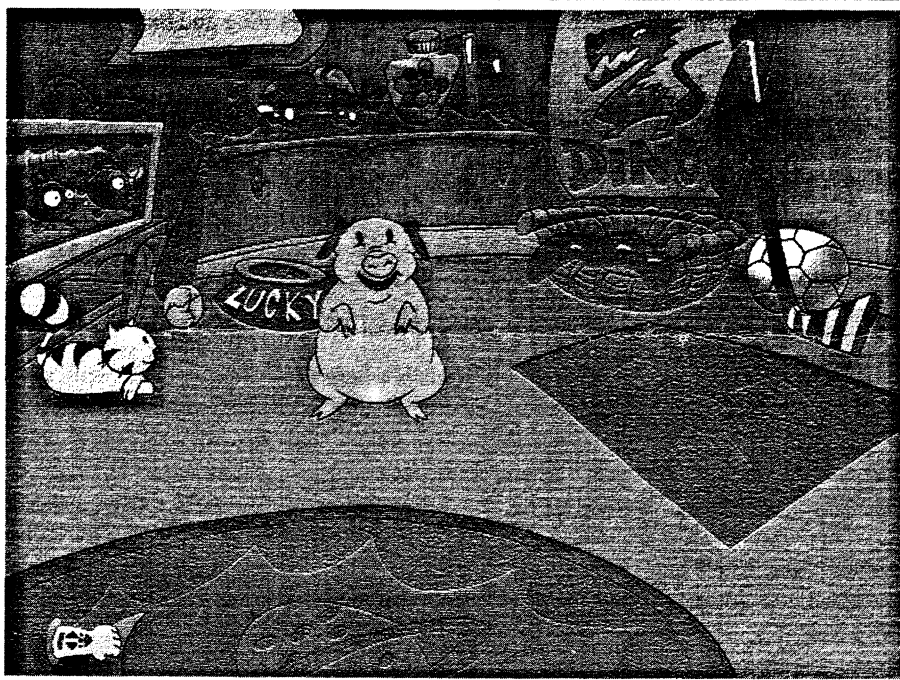
FIGURE 2:   Lucky's Corner

task accomplished, something may occur as a surprise for a job well done. If Zip does not do well, Alien catchers will pop out of the ground and lock him in a cage. In order not to be caught, Zip must work hard and learn the language. The evaluation is not a paper-and-pencil test but a colorful quiz show where Zip actually competes with other contestants.

Another feature of the program is its use of user control and program control. The user feels that he or she is in control because decisions are often made by users. In fact, the program is constantly monitoring the user's work and progress. The learner cannot go to the next mode until completing the previous one. He or she is encouraged to do one thing and discouraged to do another by receiving rewards or punishments. If the survival meter indicates a need to raise the work level, the learner is encouraged to work at the Yard Sale. Not doing so results in certain consequences.

Discussion

Balance of entertainment and education

A simulation is different from other language learning programs. The biggest strength of a simulation is its simulation of real life, which makes the language learn-

ing an interesting experience. Because real-life communication is important in language learning, simulation is a good tool to use. However, we should also realize its limitations. When we first started designing this program, we intended to cover a normal English curriculum—basic words and phrases, common sentence patterns, and so forth. It turned out to be difficult. With a game design like that of ZIP & TERRY, it is not easy to cover everything in a whole curriculum.

The program was intended to be an entertaining piece of language learning software. The educational value is important, and so is the entertainment value. Therefore, it is essential to keep a balance between entertainment and educational value. Limited by the space of the CD as well as the budget, we decided to work mainly with three major environments—Living Room, Kitchen, and Yard Sale—which cover some essentials of the curriculum, but not all. It is also realized that the program may lose entertainment value if it is too long. As we have to keep scores and give rewards/ punishments, there are limits as to what kind of learning activities we can give. At the same time, learning objectives limit how much entertainment we can provide. So how to combine entertainment with education in an optimal way was an issue we struggled with, and we believe it is the problem all the designers of educational simulations are facing and the question that requires many studies.

ESL pronunciation and speech recognition accuracy

Speech recognition technology has become much more accurate in recent years. ZIP & TERRY uses the latest in ASR technology. However, the recognition is still not 100% accurate. Our users are ESL students whose pronunciation is often different from that of native speakers. A word or a sentence can be rejected even though an ESL user's pronunciation is acceptable.

There are two kinds of ASR—Speaker-Dependent ASR and Speaker-Independent ASR. Speaker-Dependent ASR tries to adapt to a user's pronunciation and accent. As a result, the more often the user uses it, the more accurate the recognition becomes, or in other words, the fewer mistakes ASR will make. We decided not to use this kind of speech recognition for the program because we thought it would not help to improve learners' pronunciation, as it tended to accept their accent and bad pronunciation. The ASR we chose to use is Speaker-Independent. If the user's pronunciation deviates from the standard too far, it will be rejected. For this kind of ASR, the acceptance level can be adjusted. If the level is set too high, users' pronunciations will be more likely to be rejected. If the level is set too low, it will accept anything as correct. Again, here we tried to find a compromise—set the level low enough to accept ESL pronunciation but high enough to eliminate recognition errors. This was done through testing with young ESL learners.

One question to be discussed is if we can use Speaker-Dependent ASR in language learning simulation. Simulation is used for communicative activities. In this case, the purpose is often not to correct the pronunciation but to keep the conversation going. It is more important for the program to understand the learner. If the purpose is to understand the learner better, there is nothing wrong with ASR adapting to learners' accents.

With the advance of computer technology, especially ASR, simulation has become a more effective tool for language learning. ZIP & TERRY represented our attempt at developing an advanced simulation with the most recent technology. The development of this program provided us with an opportunity to explore new ways to design simulations for language learning. We are collecting feedback from our users. The experience with ZIP & TERRY will help us in designing our future language learning programs.

References

Randel, J. M., Morris, B. A., Wetzel, C. D., & Whitehill, B. V. (1992). The effectiveness of games for educational purposes: A review of recent research. *Simulation & Gaming, 23*, 261-276.

Schank, R. C., & Cleary, C. (1995). *Engines for education.* Hillsdale, NJ: Lawrence Erlbaum.

Rong-Chang Li received his master's degree in English as a second language and PhD in educational technology from the University of Illinois at Urbana-Champaign. His major research interest is in Web-related language. He has published a dozen articles on how to use the Web for language learning (www.rong-chang.com.

David Topolewski has been involved in developing spoken English software technology solutions for Asian markets for several years. Prior to that he managed video game development. He earned his BA at Yale University and his MBA from the Wharton School at the University of Pennsylvania.

ADDRESSES: RL: English Xchange, Inc. 1545 Wilshire Blvd, Suite 511, Los Angeles CA 90017, USA; telephone: +1 626-574-9883; fax: +1 626-574-9883; e-mail: rli8114@yahoo.com. DT: English Xchange, Inc. 1545 Wilshire Blvd, Suite 511, Los Angeles, CA 90017, USA; e-mail:davidtopolewski@cs.com.